1997

Autocorrelation techniques for soft photogrammetry

Wu Yao
Iowa State University

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Autocorrelation techniques for soft photogrammetry

by

Wu Yao

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Civil Engineering (Geometronics)
Major Professor: K. Jeyapalan

Iowa State University
Ames, Iowa
1997

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Wu Yao

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

I, J: Pixel coordinate system for the photographs before interior orientation

$size\_I, size\_J$: Photo size of the photograph before interior orientation

$g(I, J)$: Gray value of pixel (I, J) on the photograph before interior orientation

I_p, J_p: Pixel coordinate system for the photographs after interior orientation

x_p, y_p: Photo coordinate system

x_{0p}, y_{0p}: Initial photo coordinates of a photograph

C_{xp}, C_{yp}: Pixel sizes of photograph after interior orientation in x and y directions

$size\_x_p, size\_y_p$: Photo size of the photograph in photo coordinate system

$g_p(I_p, J_p)$: Gray value of pixel (I_p, J_p) on the photograph after interior orientation

I_c, J_c: Pixel coordinate system related to the camera coordinate

x_c, y_c: Camera coordinate system

x_{0c}, y_{0c}: Initial camera coordinates of a photograph

C_{xc}, C_{yc}: Pixel sizes of photograph in x and y directions in camera coordinate

$size\_x_c, size\_y_c$: Photo size of the photograph in camera coordinate system

$g_c(I_c, J_c)$: Gray value of pixel (I_c, J_c) on the photograph in the camera coordinate system

I_r, J_r: Pixel coordinate system related to rectified coordinate

x_r, y_r, z_r : Rectified coordinate system

x_{0r}, y_{0r}: Initial rectified coordinates of a photograph

C_{xr}, C_{yr}: Pixel sizes of photograph after rectification in x and y directions

$size\_x_r, size\_y_r$: Photo size of the photograph after rectification

$g_r(I_r, J_r)$: Gray value of pixel (I_r, J_r) on the photograph after rectification

I_s, J_s: Pixel coordinate system related to the equal scale coordinate

x_s, y_s, z_s: Equal scale coordinate system

x_{0s}, y_{0s}: Initial equal scale coordinates of a photograph

C_{xs}, C_{ys}: Pixel sizes of photograph after in x and y directions in equal scale coordinate

$size\_x_s$, $size\_y_s$: Photo size of the photograph in the coordinate after equal scale transformation

$g_s(I_s, J_s)$: Gray value of pixel (I_s, J_s) on the photograph after equalizing the scale in a stereo pair

I_e, J_e: Pixel coordinate system related to epipolar coordinate

x_e, y_e, z_e: Epipolar coordinate system

x_{0e}, y_{0e}: Initial epipolar coordinates of a photograph

C_{xe}, C_{ye}: Pixel sizes of photograph after epipolar rectification in x and y directions

$size\_x_e$, $size\_y_e$: Photo size of the photograph in epipolar coordinate system

$g_e(I_e, J_e)$: Gray value of pixel (I_p, J_p) on the photograph after epipolar rectification

x_{min}, x_{max}, y_{min}, y_{max} : Limits of in x and y axes of a digital image

X_0, Y_0, Z_0: Location of camera in ground coordinate system

X, Y, Z: Ground coordinate system

a, c, b, d, x_0, y_0: 6 parameters for the interior orientation transformation

f: Focal length of camera or principal distance

ω, ϕ, κ: Three independent angles used to determine the relationship between ground and camera coordinate

θ: Angle rotated around z_s axis to transform from equal scale coordinate to epipolar coordinate

$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$: Rotation matrix defines the direction relationship between ground coordinate system and the camera coordinate system

M: Matrix used in the computation of rectification of single photograph

S: Matrix used in the epipolar rectification

$\mathbf{g}(I, J)$: pixel value on pixel (I, J)

$\mathbf{g}_o(I, J)$: pixel value on pixel (I, J) after interior orientation

C_x, C_y: pixel sizes in x and y directions respectively

$x_0^{'}, y_0^{'}$: Photo coordinates of the upper-left corner pixel on a digital photograph

x_0, y_0 : Camera coordinates of the upper-left corner pixel on a digital photograph

ABSTRACT

In this thesis research is carried out on image processing, image matching searching strategies, feature type and image matching, and optimal window size in image matching. To make comparisons, the soft photogrammetry package SoftPlotter is used. Two aerial photographs from the Iowa State University campus high flight 94 are scanned into digital format. In order to create a stereo model from them, interior orientation, single photograph rectification and stereo rectification are done.

Two new image matching methods, multi-method image matching (MMIM) and unsquare window image matching are developed and compared. MMIM is used to determine the optimal window size in image matching. Twenty four check points from four different types of ground features are used for checking the results from image matching. Comparison between these four types of ground feature shows that the methods developed here improve the speed and the precision of image matching.

A process called direct transformation is described and compared with the multiple steps in image processing The results from image processing are consistent with those from SoftPlotter. A modified LAN image header is developed and used to store the information about the stereo model and image matching. A comparison is also made between cross correlation image matching (CCIM), least difference image matching (LDIM) and least square image matching (LSIM). The quality of image matching in relation to ground features are compared using two methods developed in this study, the coefficient surface for CCIM and

the difference surface for LDIM. To reduce the amount of computation in image matching, the best-track searching algorithm, developed in this research, is used instead of the whole range searching algorithm.

CHAPTER 1 INTRODUCTION

Photogrammetry is the science of obtaining the location, shape and size of objects by measuring them using aerial photographs. It has many applications in areas such as civil engineering, mapping and medicine. Photogrammetry is described under different headings such as topographic photogrammetry and non-topographic photogrammetry, and close-range photogrammetry. When the photogrammetric work is done by optical and mechanical instruments, it is known as analogue photogrammetry, and when combined with digital computers, as analytical photogrammetry. The latest development in photogrammetry, which is called soft photogrammetry, is mainly based on the computer and related digital techniques. The main differences between the conventional photogrammetry and the soft photogrammetry are that in the latter, computers do most of the work, such as rectifying the aerial photographs, the interior and exterior orientation, creating maps etc. automatically or interactively using only digital data processing techniques, without any other optical or mechanical parts. In addition to the human operator only digital data and digital processing devices like computers are involved in soft photogrammetry. With the computer, operators can do the photogrammetric work much faster and perhaps even better than with the conventional analogue or analytical photogrammetric instruments.

1.1 Development of Photogrammetry

As in the development of other computer-based technologies, the field of photogrammetry underwent three stages of development, from analogue to analytical and then to digital viz., soft photogrammetry.

1.1.1 General Procedures in Photogrammetry

Although so many different types of photogrammetric instruments have been or are being used, the data processing procedures involved are similar. A particular photogrammetric instrument is used in each step of the procedures. These procedures can be applied no matter what kind of imagery is used. The procedures are:

1. Interior orientation

2. Exterior orientation or triangulation

3. Rectification

4. Compilation of different types of maps

The first three procedures involve the pre-processing of the aerial photographs. Interior orientation registers the photographs into the instrumental coordinates like photo coordinate and removes the image deformation caused by the camera imaging system by using the camera calibration data. The exterior orientation or triangulation is used to determine the orientation of the photographs. Rectification is used to remove the errors brought by the tilt of the camera when the photograph was taken.

The fourth procedure is for the purpose of making maps. Different types of maps, such as topographic maps or contour maps can be made.

Three types of photogrammetry will be discussed in the following sections: analogue, analytical and digital photogrammetry.

1.1.2 Analogue and Analytical Photogrammetry

1.1.2.1 Analogue Photogrammetry

Many of the analogue photogrammetric instruments such as analogue stereo plotters were designed before 1950 [Blachut and Burkhardt, 1989]. At that time, only optical and mechanical devices were used in photogrammetric data processing. After the invention of analogue and digital computers, computer techniques were also used in sampling and processing the data on the analogue photogrammetric instruments. All of the modern analogue stereo plotters use computers for data digitizing and processing.

1.1.2.2 Analytical Photogrammetry

Since the 1950s analytical photogrammetry has had more wide applications [D. Frederick, 1964; Ghosh, 1979 , 1988; Church]. In analytical photogrammetry, the solution of problems are by mathematical computations. The basic theory of analytical photogrammetry had already been developed long before the design of the analytical photogrammetric instruments. In analytical photogrammetry, the solution of problems are by mathematical computation, so it was only after the development of computer technology that analytical photogrammetry came into wide use.

The first analytical stereo plotter was introduced by Dr. U. V. Helava in 1957 [Lanckton, 1969; Helava, 1958, 1960; Hohnson, 1961]. In analytical photogrammetry, computer techniques are used in interior orientation, triangulation, and to manage the data in map compilation. Spatial analysis techniques are applied to analytical photogrammetry so the operator can pick up the ground features like break lines which can be stored and processed easily by the computers. Compared to analogue photogrammetry, analytical photogrammetry techniques are more efficient and accurate. But they also need operators for map compilation.

In analogue, as well as in analytical photogrammetry, optical and mechanical systems are still used to display and measure images. The photographs are in analogue form in both types of photogrammetry.

1.1.3 Soft Photogrammetry

Soft photogrammetry is also called Digital Photogrammetry or Softcopy Photogrammetry.

Soft photogrammetry is totally different from the analogue and analytical photogrammetry in the method of image displaying and processing. Only digital image data are used and processed in soft photogrammetry. In soft photogrammetry the triangulation, rectification and compilation of maps mentioned before can be automatically performed.

1.1.3.1 Development of Soft Photogrammetry

The main goals of soft photogrammetry are (1) the automation of map creation from imageries and (2) the digitalized system. The capacity of computers and digital electronic devices for image and data storing and processing provide a very powerful tool for photogrammetry. In the beginning, analogue computers combined with optical instruments were used in soft photogrammetry. In this early soft photogrammetry system, electronic correlator and microdensitometer were used to digitize and match the images in a stereo pair [Hobrough, 1959; Bertram, 1963; Helava, 1966]. But soon different types of computerized photogrammetric systems and mathematical algorithms with analytical plotters were used to achieve a high level of automation in photogrammetric data processing [Scarano, 1976; Helava, 1978; P. Alfred, 1984; Masry, 1974]. In these systems, additional electronic systems, such as a microdensitometer, are added to the analytical plotter to convert the image to digital signals for one small area at a time, and an analogue or digital computer is used to do the

image correlation. Many of these old systems were on-line systems. Now all imagery data are stored, manipulated and displayed by computers. Therefore, the performance in soft photogrammetry largely depends on the following factors:

For the hardware:

- the storage size of the computer disks
- the speed of the CPU and data accessing to the storage disks
- the capacity of the computer memory
- the quality of the computer monitor for displaying images
- the precision of the digitizing or scanning device

For the software:

- the quality of the triangulation software
- the quality of the matching algorithms
- the speed of the matching algorithms
- the speed and quality of image compression/decompression
- Scale of the imageries
- Ground variance of the area of interest
- Characteristics of ground features

Since digital files can be very large(> 100 MB) in high precision images, and the data processing takes time, CPU speed in data processing is a very important factor. Therefore, soft photogrammetry systems are mainly based on workstations or a main frame computer system. With the parallel computation technique, a PC computer system with more than one CPU can work as fast as workstations. For some low resolution work a high speed PC may be sufficient. Since Windows NT operating systems can works on different types of CPU or processors, more and more soft photogrammetric system are based on Windows NT operating systems.

An image matching algorithm is the most important consideration in the software. It determines the precision and speed of map compilation. Since the software used to process the imageries in soft photogrammetry is the only interface that the operators can use to

control the procedure of processing, user friendliness, efficiency and flexibility of the software are also very important considerations.

An image compression/decompression technique is needed because of the quantity of data. An image compression technique can save a lot of disk storage space. Since image compression/decompression will take time, its speed is important to the data processing. Hardware can be used to speed up the image compression/decompression process.

Scale of the imageries, ground variance, and characteristics of the ground features also affect the precision of the soft photogrammetric process. The larger the scale and ground variance, the less similar the images in a stereo pair because of more obstructions from the height variance.

There are many factors in soft photogrammetry that will affect the output.

With about 30 years of development in soft photogrammetry, it is already better than analytical photogrammetry.

1.1.3.2 General Procedures of Data Processing in Soft Photogrammetry

A soft photogrammetry system consisting of only a computer system, a digitizing device (Scanners) and software, is totally different from the conventional photogrammetry system. Since all the data are in digital format, the data processing is also different from that in conventional photogrammetry. There are also different imagery sources such as aerial photographs, satellite imageries and video imageries. The general steps to process imageries are:

1. Digitize them by scanning if they are not in digital format
2. Perform interior orientation
3. Perform triangulation and relative orientation
4. Rectify the digital imageries
5. Perform epipolar or stereo rectification
6. Pre-process the image data
7. Perform image matching
8. Post-process the image matching output
9. Create DEMs, orthophotos and/or maps like contour maps

Image resampling needs to be applied in the above procedures before the image matching. Since image resampling will change the values of the pixels in an image and will affect the image matching, it would be better to have few image resamplings. Steps 2, 3, 4, 5 can be performed together thereby reducing image resampling to only once.

Different image processing techniques can be used in the image pre-processing to improve them before image matching.

Post-processing gives the operator a chance to interactively edit the image matching results.

1.1.3.3 Advantages of Soft Photogrammetry

Automation of the photogrammetric process is the biggest advantage of soft photogrammetry. It can produce maps without or with only little interaction with operators. Once all the parameters are set, the operator can just leave the computer program running and attend to other work. Since all image data are in digital format and stored on small disks, and the only instrument needed is the computer, the soft photogrammetry system unlike the conventional photogrammetry is simple and easy to manipulate. Once the image data and output maps in soft photogrammetry are created and stored in computers, they can be retrieved quickly and easily any time they are needed. Since all the imageries are displayed on a computer screen, they are more comfortably viewed than in conventional photogrammetric instruments.

There is flexibility in manipulating the display system in soft photogrammetry which is probably another of its advantages. A digital image processing technique provides a powerful tool for enhancing the displaying images for the operators. Theoretically, the brightness and contrast of the images displayed on screen can be used to show better the details of both bright and dark areas of the imagery. This is done by using image processing techniques such as changing the histograms. Image filtering and restoration can also be used to preprocess the imageries for better image matching results [M. Ehlers, 1982; Dowman, 1984; J. O. Fallvik, 1986]. This facility for image filtering and restoration makes soft photogrammetry a more flexible and powerful system unlike conventional photogrammetry where only changes in brightness of the images can be done to obtain better matching results. Unlike the limited

number of choices of magnifications in the optical systems used in conventional photogrammetry, image processing techniques in soft photogrammetry can also supply more magnifications for better measurement or viewing.

Sometimes it is difficult for operators to identify or measure points in an obstructed area on the two photographs in a stereo model. In soft photogrammetry more than two images can be used in a stereo model. So the obstructed areas can be seen on the third or forth image, and the reconstruction of all faces of an object is possible [A. Gruen, 1987; T. Schenk, 1992].

Scale is not a problem when the imageries in the stereo model have approximately the same scales. But, if in the application of close-range photogrammetry, the scales of the imageries in the stereo model are much different from each other, soft photogrammetry can make the imageries in the same stereo model have the same scale by image resampling during image processing [F. Schneider, 1992]. Different types of image matching algorithms can be applied for different types of images to improve the image matching precision.

In soft photogrammetry, the aerial photographs are first converted to digital forms by scanning or digitizing. Then computers are used to display and process the digitized photographs. Once the necessary parameters or measurements are set or done, computers are used to rectify, mosaic and subdivide the digital aerial photographs without operator interruption; and computer techniques are applied to do the relative and absolute orientation of the stereo pair and aerotriangulation, and to produce digital three dimensional model and orthophotographs automatically. The output of soft photogrammetry can be Digital Elevation Models (DEMs), which represent the ground surface in digital format, and digital orthophotographs. These digital format outputs can be delivered easily by different types of disks or tapes, and retrieved, displayed and printed out as hard copy at any time. Photogrammetry produces a hard copy which is the conventional map while soft photogrammetry produces digital computer files known as the soft copy.

1.1.3.4 Disadvantages of Soft Photogrammetry

Some of the obvious disadvantages of soft photogrammetry are: high performance requirement for the computer; large storage space required for huge digital imagery files and output digital maps; high resolution screen and printing devices for better looking output

maps; expensive hardware and software; high expense needed for software maintenance; and expensive scanning devices. The reliability of the storage is also a big problem. In order to avoid losing data, all the digital data must be backed up which will take even more storage space.

In the orientation of a stereo pair, creation of DEMs and digital orthophotographs automatically, image matching or image registration is the key technique. There are various types of mathematical algorithms used in the image matching each with its unique set of advantages and disadvantages.

1.2 Image Matching and Related Image Processing Algorithms Used in Soft Photogrammetry

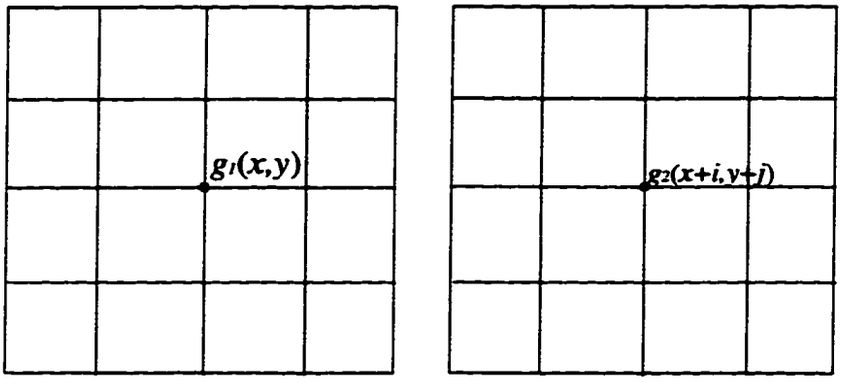
Image matching can be performed in many different ways. These image matching methods can be grouped into three types [Ghaffary, 1985]: area-based image matching; feature-based image matching; or hybrid-based image matching which is a combination of these two methods. As shown in Figure 1.1, the digital image is made up of thousands of pixels arranged in a certain format such as the grid points in Figure 1.1. The center of each pixel is used to represent the location of that pixel. Let $g_1(x,y), g_2(x+i,y+j)$ represent the gray level of the pixels located on (x, y) on the left image and $(x+i, y+j)$ on the right image in a stereo pair. The purpose of image matching is to find out the values of i, j so that the pixels $g_1(x,y), g_2(x+i,y+j)$ are the images of one identical ground point on the digital photographs.

Thus image matching can be done with area-based or feature-based image matching algorithms, it can be done in spatial or spectral space, and it can be done in the image or object space. In order to obtain a better image contrast or higher signal-to-noise ratio (SNR), some types of image processing algorithms can be used as pre-processing for image matching.

1.2.1 Area-based Image Matching (AIM)

In AIM, the gray level similarity of the pixels in a certain area on the photograph, called the mask or window, provides the information for image matching. The size of the window determines the number of pixels involved in the matching, and the location of the window

determines which pixels will be used in the matching. Since all the pixels within the search limits will be used in the comparison, this method can have high matching precision. The disadvantage of the AIM is that it may be time consuming because of the large amount of computation within the large search limits. Some different algorithms are used in the AIM.



a. Left Image b. Right Image

Figure 1.1 Digital Representative of Image

1.2.1.1 Cross Correlation Image Matching (CCIM)

Since each pixel in the window can be considered as one element in a two-dimensional matrix, the similarity between the pixels in the two windows can be computed by the two-dimensional correlation method that has been developed in statistics. The normalized cross correlation equation is [Quam and Hannah, 1974]

$$C(i,j) = \frac{\sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_1(x,y) \bullet g_2(x+i,y+j)}{\left[\sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_1^2(x,y) \bullet \sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_2^2(x+i,y+j)\right]^{1/2}} \qquad 1.1$$

Where the values of $x_0, x_1; y_0, y_1$ define the size of the window used in the image matching. C(i, j) is the correlation coefficient. The matched position determines the largest correlation coefficient C(i, j).

1.2.1.2 Least Square Sum Image Matching (LSSIM)

The similarity of the windows in the stereo pair can also be determined by the Euclidean distance between the pixels in the windows on the two images as

$$d(i,j)=\left[\sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1}\left[g_1(x,y)-g_2(x+i,y+j)\right]^2\right]^{1/2} \qquad 1.2$$

The position with the minimum value of d(i, j) gives the matched position. Radiometric correction can be added to the above equation for better results.

1.2.1.3 Least Difference Image Matching (LDIM)

Sum of the absolute values of the difference between the corresponding pixels in the two windows can also be used to perform the image matching. The equation for LDIM is:

$$d(i,j)=\sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1}\left|g_1(x,y)-g_2(x+i,y+j)\right| \qquad 1.3$$

The matched location is that with the least value of $d(i, j)$. Again the radiometric correction should be added to the above equation.

Among the above AIM methods, cross correlation has the highest matching precision.

1.2.2 Least Square Image Matching (LSIM)

Basically, LSIM is area-based image matching since it transfers the geometric and radiometric information about the pixels in the window from one image to the other in the stereo model. Geometric and radiometric transformation are both considered in the LSIM. Wrobel [1991] has a complete summary on the development of the LSIM. In early 1980s, the LSIM was applied by Ackerman and his team in Stuttgart [Ackerman, 1983a, 1984]. The basic principle of this method can be explained in the one dimensional case. Let the gray level matrices for the left and right photographs be $\bar{g}_1$ and $\bar{g}_2$ and assume they are superimposed on the same pixel coordinate. Then, we have

$$\begin{aligned}\bar{g}_1(x_i)&=g_1(x_i)+n_1(x_i)\\ \bar{g}_2(x_i)&=g_2(x_i)+n_2(x_i)=h_1\times g_1(a\times x_i+b)+h_0+n_2(x_i)\end{aligned} \qquad 1.4$$

Where a,b are the geometric transformation parameters; h_0,h_1 are the radiometric transformation parameters; $n_1(x_i),n_2(x_i)$ are the noise signals. The purpose of the LSIM is to

estimate the unknown parameters a,b,h_0,h_1 by the least square method which minimizes the difference between these two images at the matched points within the window. The difference is

$$\begin{aligned}\overline{\Delta g}(x_i) &= \bar{g}_2(x_i) - \bar{g}_1(x_i) = g_2(x_i) - g_1(x_i) + n_2(x_i) - n_1(x_i) \\ &= h_1 \times g_1(a \times x_i + b) + h_0 - g_1(x_i) - v(x_i)\end{aligned} \qquad 1.5$$

For the initial approximation of $b=0, a=1, h_1=1, h_0=0$, we have

$$g_2(x_i) = g_1(x_i) + \dot{g}_1(x_i)db + x_i\dot{g}_1(x_i)da + dh_0 + g_1(x_i)dh_1 \qquad 1.6$$

Substituting equation 1.6 in equation 1.5, we have

$$\overline{\Delta g}(x_i) + v(x_i) = \dot{g}_1(x_i)db + x_i\dot{g}_1(x_i)da + dh_0 + g_1(x_i)dh_1 \qquad 1.7$$

In the equation 1.7, the four unknown parameters are db,da,dh_0,dh_1. For each pixel in the interested area or window, we can construct one equation for the four unknown parameters. To solve these equations we can find the least square estimates of the four unknown parameters.

Because of the existence of the geometric and radiometric transformation, there is a resampling problem in the LSIM. From equation 1.7 we know, for a pixel position x_i, the value of $ax_i + b$ may not be an integer. The bilinear interpolation method is used and the value of $\dot{g}_1(x_i)$ in the equation 1.7 determined by the surrounding pixels. For the two dimensional case, by the same principle, we have the equations as:

$$\begin{aligned}\bar{g}_1(x_i,y_j) &= g_1(x_i,y_j) + n_1(x_i,y_i) \\ \bar{g}_2(x_i,y_j) &= g_2(x_i,y_j) + n_2(x_i,y_i) \\ &= h_0 g_1(a_0 + a_1x_i + a_2y_j, b_0 + b_1x_i + b_2y_j) + h_1 + n_2(x_i,y_j)\end{aligned} \qquad 1.8$$

The geometric transformations used in the equation 1.8 are Affine transformations. The radiometric transformations used are the same parameters as those used for the one-dimensional case. Then the total number of unknown parameters used in equation 1.8 are eight. Thus, we can write the least square equation for the first iteration as:

$$\begin{aligned}\overline{\Delta g}(x_i,y_j) + v(x_i,y_j) &= \dot{g}_x(x_i,y_j)da_0 + x_i\dot{g}_x(x_i,y_j)da_1 + y_j\dot{g}_x(x_i,y_j)da_2 \\ &+ \dot{g}_y(x_i,y_j)db_0 + x_i\dot{g}_y(x_i,y_j)db_1 + y_i\dot{g}_y(x_i,y_j)db_2 + dh_0 + g_1(x_i,y_j)dh_1\end{aligned} \qquad 1.9$$

As in the one dimensional case, resampling is also needed in the two dimensional case. By the iterative method, the solution of the least square equation 1.9 can be found.

The advantage of the LSIM is the high precision you can obtain from 1/180 to 1/10 pixel , according to the quality of the aerial photographs. But from the equations 1.6 and 1.9, the convergence of the least square equation strongly depends on the initial approximation and the feature of the images in the stereo pair. The more iteration means that more time will be spent in the computation. Therefore it may be time consuming for matching a whole stereo pair.

See [Helava, 1987; D. Rosenholm, 1987a, 1987b and 1987c] for more about LSIM.

1.2.3 Feature-based Image Matching (FIM)

In FIM the feature similarities between the images in the stereo pair are used in the image matching. There is no matching window used in the FIM. Some of the ground features used in the FIM are:

- Location of the edges, boundaries, points, center of mass, etc.
- Orientation of the edges and boundaries
- Corners and Intersections
- Shape similarities of the edges and boundaries
- Colors similarities in the color photographs
- Changes of gray level around the features

Since only these features are extracted and compared, this method will take less processing time. But the disadvantage of this method is that the matching precision is not as high as in area-based image matching. It can be used in image pre-matching and some process like automatic relative orientation.

The first step in the FIM is to extract the useful image feature. According to Ghaffary [1985], the FIM algorithms can be divided into three groups: Edge-following approach; Feature clustering approach; Symbolic matching.

Edge-following approach is the method often used. In this method, the edges are detected. The purpose of the edge detection is to extract the edges from the images. There are various edge detectors used to detect the local edges and boundaries by using the zero-crossing, and image enhancing methods [References 30 - 34 in Ghaffary, 1985; Reference 7 in Greenfeld;

Förstner, 1987 and 1986; J. Greenfeld, 1989; A. Huertas, 1986; D. Li, 1994; T. Luhman, 1986; J. Piechel, 1986; Z. C. Qiu, 1992; J. Zong, 1992]. The detected edges are then used to form line segments. These line segments are of simple form such as polygon vertices of edge pixels [References 37 - 51 in B. K. Ghaffary, 1985. Greenfeld, 1988].

1.2.4 Hybrid Image Matching (HIM): Combination of Area-based and Feature-based Image Matching

Combination of area-based and feature-based image matching can take advantage of both area-based image matching and feature-based image matching. The AIM has a high matching accuracy. In case of LSIM the matching results highly depend on the quality of the initial points. On the other hand, FIM has lower matching accuracy but high matching speed. So the FIM can be used to provide the good initial positions for AIM.

Some examples of the HIM are given by Helava [1987], M. Li [1990]. The feature information is extracted and checked by some algorithms. The results are then used as the initial approximation of the LSIM. A four parameters transformation is used. Two of the parameters are for the geometric shift, and the other two are for the radiometric transformation. The estimated precision for the text photographs for stereo parallax measurement is 0.06 pixels.

1.2.5 Object Space Image Matching

Image matching can be performed in image or object space. When matching in object space, the location in the real world coordinate system is considered and converted back to the image space. The advantage of the object space image matching is that the images in the stereo pair do not have to be rectified into an epipolar coordinate system and single photograph rectification is not needed. Only the triangulation results of the photographs are used in the image matching [M. Benard, 1986]. The image matching algorithms which can be used in image space matching can also be used in the object space matching.

Object space image least square methods are also being developed [Helava, 1988; Wrobel, 1987; Lo, 1992]. These methods consider the ground surface consisting of small elements or facets and do the matching directly in the object space. In the object space image least square

methods, the elevations of the ground elements at the gray level on the photographs are considered and the least square equations are formed.

Vertical Line Locus [M. Benard, 1986] uses the image pixels from location in the object space to do the image matching. Only the elevation of the location in the object space is changed during the image matching. For each location in object space, the height of that location is changed until the required image matching is reached. Dual space image matching is also being researched [Y. N. Zhang, 1992].

1.2.6 Spatial and Spectral Image Matching

Image matching can be performed in spectral space. Fourier transformation is used to transform the image from spatial space to spectral space [R. A. Schowengerdt, 1983; Anuta, 1970]. Assume the correlation between functions $g_1(x, y)$ and $g_2(x, y)$ is expressed as:

$$c(i, j) = g_1(x, y) * g_2(x, y) \qquad 1.10$$

then the Fourier transformation of the above equation is:

$$C(v_i, v_j) = G_1(v_x, v_y) G_2^*(v_x, v_y) \qquad 1.11$$

Where C, G_1, G_2 are the Fourier Transformation of c, g_1, g_2 respectively; G_2^* is the complex conjugate of G_2; $v_i, v_j; v_x, v_y$ are the variables in the spectral space.

Fourier transformation converts the correlation computation in the spatial space to product in the spectral space. Fast Fourier Transformation (FFT) are used to facilitate the process. FFT is generally used for images less than 512 by 512 pixels. For more information see the references space [R. A. Schowengerdt, 1983; Anuta, 1970].

1.2.7 Optimal Window Size in Area-based Image Matching

Window size is an important factor affecting the speed and precision of image matching. Dan Rosenholm [1987a] using LSIM and square window size 12x12, 16x16, 20x20, 30x30, 40x40 and 50x50 with Affine and without Affine transformation to check the precision, concluded the optimal window size to be between 20x20 and 30x30 pixels. M. Li [1990] obtained almost the same precision curve, but the optimal window sizes range is a little larger.

1.2.8 Image Pre-processing in Image Matching

Different types of image processing, such as filtering [Dowman, 1984; M. Ehlers, 1982; S. B. Abramson, 1993; Adamos, 1992; Z. X. Zhang, 1992], and image restoration [J. O. Fallvik, 1986] can be applied before or during performing image matching. These image processing methods can improve the quality of the imagery for better image matching results in AIM, or obtain information from the images for the image matching in FIM.

1.2.9 Multi-Image Matching

More than two images can be used in image matching [Rosenholm , 1987c; A. W. Gruen, 1987; T. Schenk, 1992]. The advantage of multi-image matching is that more than one stereo pair over the same area can be used in the matching. Therefore the obstructed area on one stereo pair might be seen on other pairs. It is also helpful to use multi-images in generating orthophotos.

1.2.10 Factors That Affect Image Matching

Some of the factors that may affect the precision of image matching are:

- image texture (gradients of gray-value functions)
- pixel size
- geometric precision of the pixels
- window size
- window shape
- image preprocessing
- signal to noise ratio (SNR)

1.2.11 Image Pyramid in Image Matching

Image pyramid or tiling technique is used in image matching [El-Hakim, 1989; Kaiser, 1992; S. Hattori, 1986]. In the image pyramid the images in the stereo pair are resampled into different lower resolutions of power of 2 such as 2, 4, 8, 16. The image on the top of the pyramid has the lowest resolution or the largest pixel size. Image matching starts at the top tile of the image pyramid, then goes to the next one on the pyramid and the matching results from the previous one will be improved. This important technique used in image matching. It

can save a lot of processing time and improve the matching results. The disadvantage of this method is that it will use more storage space for the extra images.

1.2.12 Automatic Relative Orientation

Since automatic image matching can be used to transfer the points from one image to another in the stereo pairs, automatic relative orientation is also available [J. S. Greenfeld, 1990; Li M., 1990; J. Q. Zhang, 1992]. FIM is better for the automatic relative orientation since the geometric constraints for the stereo pair are not established before relative orientation.

1.2.13 Close-range Applications

Close-range application is different from the aerial photograph application in photogrammetry. Large scale and angles may be involved in the close-range application. soft photogrammetry can be used in the close-range applications [G. P. He, 1992].

1.2.14 Different Scale Image Matching

Soft photogrammetry can used to match images with different geometric scales automatically [F. Schneider, 1992]. Image processing techniques make it easy to change the images with different scale into the same scale, either in image rectification or matching.

1.2.15 Other Image Matching Methods Used in Soft Photogrammetry

There are some other methods used in image matching, such as multi-criterion matching [Z. J. Lin, 1986], multi-template matching [Kölbl, 1987], phase shift matching [J. Stokes, 1986]. Other new technologies such as artificial neural networks [L. Gong, 1992] are also being used in soft photogrammetry.

1.3 Software in Soft Photogrammetry

Since computer software in soft photogrammetry is used to implement all the image matching algorithms discussed above, the software packages are the most important part in soft photogrammetry. Therefore the quality of the software in soft photogrammetry represents the level of soft photogrammetry.

Two soft photogrammetry software packages, SoftPlotter [Vision International] and Socet Set [HELAVA Associates Inc.], will be compared in this section. The current development of software in soft photogrammetry is presented through this comparison.

SoftPlotter is a soft photogrammetry software package from *Vision International*, and Socet Set is a soft photogrammetry software package from *Heleva Inc.*. Both software packages can be used to perform what can be done by conventional photogrammetry to the digital images, and they have versions based on different types of workstations. The ones used in this research are based on a UNIX operating system on SGI workstation. SGI workstations have strong image processing ability and high quality image display which are suitable for soft photogrammetry applications.

The general advantage of SoftPlotter is its simple and friendly user interface. Everything in it is grouped into modules with certain functionality, so that the user knows how to process the images according to the information from each module. This is different from Socet Set. In Socet Set all the functions are grouped by different levels of menus in terms of the similarity of processing procedures, which makes it not very user friendly. In terms of functionality, Socet Set has more complicated and complete functions that can be used to deal with more different situations, although the SoftPlotter has sufficient accuracy for all applications.

1.3.1 Appearance and User Interface

SoftPlotter has a module menu called tool bar that shows all the available function modules on it. All the modules are arranged in the proper order of processing. Therefore, it is possible for the users to use the SoftPlotter for basic operations without or with little help from the manual.

In Socet Set, there is no such module menu. Basically it consists of three separate windows vertically arranged on the screen. The upper one has all the functions in the menu, the middle one is for displaying images, the lower one controls the image display in the middle window. For dual-head displaying, there is another window on the secondary monitor. Since all the functions in the upper window are not grouped into function modules as in the SoftPlotter, it is not as user-friendly.

The general impression is that in relation to the functions or operations the Socet Set is more complicated than SoftPlotter. Since SoftPlotter has all its functions in modules and arrange them on the tool bar in the order of processing, it is more efficient than the Socet Set. It is not difficult to add more complicated functions, if needed, to SoftPlotter to improve its processing precision or functionality for each module. Table 1.1 presents the main menu comparison between the SoftPlotter and Socet Set.

Table 1.1 Main Menu Comparison

	SoftPlotter	Socet Set
Functions in modules?	Yes	No
User friendly?	Yes	Not much
Image window always available?	No, different graphic window for different function	Yes, always the same graphic window
Saving current status when exiting?	No, it uses default project	Yes, it will come back to the same status next time
Is the main menu complicated?	No, there is only one tool bar	Yes, there are three separate windows

1.3.2 Creation of New Project

In both software packages, the first step in data processing is to create a new project. In the SoftPlotter, information like project name, projection or datum type, average height, output image format and ground spacing of the output DEM are needed. Once a new project is created, then a subdirectory under the project name is created and a certain number of subdirectories under that project subdirectory are created for storing the input and output for corresponding modules. Unlike SoftPlotter, in Socet Set the paths of images and project files are needed for the new project. Input and output information will be put in these two subdirectories. If the default path is not used, the subdirectory for the new project needs to be created before setting the path to define the new project. The coordinate system and estimate of the minimum and maximum ground elevations are needed for the definition of the new project. Table 1.2 shows a comparative account between the SoftPlotter and Socet Set when creating a new project.

Table 1.2 Creating New Project

	SoftPlotter	Socet Set
Subdirectory for each module?	Yes. This makes it easy to access the results	No, only two subdirectories for images and project
Easy to create new project?	Yes	Fine
Need ground height estimate?	Yes	Yes
Need ground spacing for DEM	Yes	No

1.3.3 Project Information Management

In SoftPlotter all the information is kept in the corresponding subdirectory for the module. For example, the information about the photographs and ground points in the current project are kept in the files in the subdirectory called the *block*. In this way it is easy to have the information you need, and monitor its status. In Socet Set, the information are kept in the files in the same subdirectory. So it is not easy to find the information if you are not familiar with the project structure of the software. On the other hand, since the SoftPlotter has different modules, and the status information are displayed when any of the modules is opened, it is easy for users to get the information they need about the current status of the module or the project.

1.3.4 Image Importing

Both software packages can use different types of digital imagery formats. Since different types of imagery format can be converted to each other, it is not a problem for them to use different formats. But the way that they import imageries is slightly different. It is interesting that Socet Set can not read TIFF images correctly with more than one strip.

In SoftPlotter the images are imported after selection of the camera. There is a camera database which must contain at least one camera before the image frames can be read. The exterior orientation parameters or the aerotriangulations can be imported from some other results like Albany. In the importing process, all the information on the imported frames are displayed in the information table.

In Socet Set, the camera calibration data are stored in different camera data files. Before the image is imported, the camera data file has to been chosen. The exterior orientation can be imported as one of the options. The imported images are arranged as separate files. It is not

easy to get the information on the imported images. There is a function that can be used to group stereo models from the imported images. Table 1.3 summarizes this comparison.

1.3.5 Camera Calibration Data

In SoftPlotter all the camera calibration data are stored in a central camera database that can be used by any project. In the *Block* tool of each project, there is another local camera database which is only valid for the current project. The calibration data in the database can be modified at any time.

Table 1.3 Comparison on Image Importing

	SoftPlotter	Socet Set
Can it import triangulation?	Yes, at any time during processing	Yes
Can it use multiple cameras?	Yes, any camera in the database	Yes
Can it modify camera data?	Yes, at any time during processing	Yes
Can it display information for all the frames?	Yes, displayed in a table	No
Easy to import?	Yes	Fine, little problem with importing TIFF images

In Socet Set the calibration data for one camera are stored in a file. There is a function that can be used to modify the camera calibration data in the camera calibration data file. Since the camera calibration data for each camera are stored in separate files, it is not easy to know the number of cameras and the types of cameras used in a project. But it is not difficult to modify the content of the camera calibration data using the function in Socet Set. It has a very good graphic interface for users to locate the fiducial point and input the photo coordinates.

Since SoftPlotter uses database to handle the camera calibration data, it is easy for its users to manage the camera data.

1.3.6 Interior Orientation

Both software packages have good graphic user interfaces to do the interior orientation.

In SoftPlotter, the basic image processing tools like magnification are in the same window as other information. But in Socet Set, there is a separate tool window for these basic image processing. The latter is better but more complicated.

In Socet Set, the fiducial points can be enabled or disabled by checking on option *yes* or *no*. In SoftPlotter, for those unused fiducials, the users do not measure them.

Both software packages provide the correction for radial distortion of the camera.

1.3.7 Ground Point Measurement

The process used in Socet Set for measuring ground points is complex for a beginner. There is a function that can be used to input all the ground points as control, tie, and check points. The points can also be added or deleted in the measurement window. Since only two images can be displayed at a time, in order to measure the points on more than two images, a new image has to be loaded after measurement is done with measuring. It has one cursor for each image. One cursor has to be locked before the measurement on the other image can be made. Although only two images can be displayed at a time, any number of images can be loaded if the points to be measured appear on them. Automatic point measurement can also be done.

It is easier to measure the points in SoftPlotter than in Socet Set. Since the ground point information is stored in a table which can be accessed from almost any where in the *Block* tool, its content can be modified very easily. The ground point information can also be imported from a text file. When measuring the ground points, up to 6 images can be displayed on the graphic window at the same time. Automatic measurement is also available for both pug and control points. Once the point is measured, the point number is marked on the displayed photograph. Another big advantage for SoftPlotter is that the frame information can be accessed from the ground point measurement window and the triangulation can be performed immediately after the ground point measurement is done. It takes little time to do triangulation and see the triangulation results after finishing the ground point measurement. The triangulation results from each iteration are displayed for checking. Table 1.4 presents

some comparisons. Since all the information about the ground points are tabulated, it is easy to check and modify the values for each point. Since up to 6 images can be displayed at the same time in SoftPlotter, sometimes this can cause memory problems such as memory overflow or segment violation. One of the reasons for this may be that some previously claimed memories are not released when they have been used. It also has some problems with color images. It is slow in loading a certain type of color images even when the images are not that big.

Table 1.4 Comparison of Ground Point Measurement

	SoftPlotter	Socet Set
User friendly?	Yes	OK
Efficient?	Yes	No
Measured points marked?	Yes	No
Automatic measurement available?	Yes, for both pug and control points	Yes
Can it import ground points?	Yes	No
Can it display more than two images?	Yes	No
Any memory managing problem?	Yes, memory error for big image(i.e. 160MB)	Not test
Easy to activate or deactivate points?	Yes	OK
Complicated?	No	Yes

1.3.8 Triangulation

There is no problem in performing triangulation in SoftPlotter. The triangulation can be performed either in ground point measurement or from the menu in *Block* tool. The results of triangulation can be viewed immediately after the triangulation is done, whether the results are good or not. All the information about the triangulation are stored in a text file which can be easily accessed. Also, the triangulation results can be dumped into a text file or sent to the printer.

The same control and point measurements used in SoftPlotter were carried out in Socet Set. The triangulation program of the Socet Set was unavailable for this research.

1.3.9 Other Functions

The results of other functions like automatic creation of stereo pair, DEM, triangulated irregular network (TIN) and orthophotos, three-dimensional digitizing are available only from SoftPlotter. All the normal photogrammetric operations can be performed automatically or manually in SoftPlotter, such as the creation of contour and planimetric maps.

1.3.10 Summary

All the operations in SoftPlotter are easy to perform. It is user-friendly, whereas the same operation in Socet Set can be complicated. There is no doubt that Socet Set is a good softcopy photogrammetry software. It has more functions than SoftPlotter and probably produces better results in some cases. From the user's point of view, however, SoftPlotter is much more user-friendly, and it can also produce high quality output maps.

From the above comparison of the two soft photogrammetry software packages, it can be established that the current soft photogrammetry software packages have the following characteristics:

- Complete functions for all the operations in photogrammetry.
- Automatic creation of DEM, TIN and orthophotos.
- Manual processing is available which is the same as the operations in conventional photogrammetry. In manual procedure the only difference from conventional photogrammetry is that digital images are used.
- User-friendly interface for easy learning and using.
- High precision image matching.

1.4 Objective of This Research

Soft photogrammetry has developed to a high level of efficiency, but improvements can be made. This research attempts to find a way to improve the image matching accuracy by using different searching strategies, image matching algorithms, and matching windows. Also, computer programs are developed for processing the images.

The aerial photographs used in this research are the high flight aerial photographs No. 1 and 2 in the Iowa State University campus project of 1994.

In order to evaluate the results, a project with the same control and ground point measurements is created in SoftPlotter and 24 points with known heights are used. These 24 check points are of four different types:

- 6 control points
- 6 points located on the ground with sharp contrast
- 6 points located in grass or tree areas
- 6 points located at corners of building roofs or top of the water tower

The heights of points other than control points are manually determined in SoftPlotter. All the matching results are evaluated by using these 24 check points.

The main purpose of this research is to complete a procedure for matching digital images from aerial photographs and to study the image matching algorithms by soft photogrammetric methods. Specifically, its objectives are:

- to obtain the pre-processing of the stereo pair. This pre-processing includes interior orientation, and single and stereo rectification. All the formulae suitable for programming are derived. A modified LAN image format suitable for storing the digital images in soft photogrammetry is used, and all the information about the stereo model are stored in the modified LAN header. This is discussed in Chapters 2, 3, 4 and 5.
- to study a new searching strategy and characteristics of CCIM and LDIM. The surfaces of the coefficient of CCIM or the difference in LDIM are displayed and used to do the analysis. This is discussed in Chapter 6.
- to study multi-method image matching and its application in determining optimal window size for different stereo pairs. Results of the 24 check points with different types of ground features are compared. This is discussed in Chapter 7.
- to study unsquare window image matching. The results of the 24 check points are obtained and compared in Chapter 8.

Through this research, a stereo pair of digital images in modified LAN format is created from the raw aerial photographs. The results for the four different types of ground points used to evaluate the methods are consistent with the results from SoftPlotter. A new searching strategy, best-track searching, is compared with whole range searching using CCIM and

LDIM. The results show that the best-track searching strategy can give the same precision as the whole range searching when the window size is large enough. Also, the surfaces of correlation coefficients from CCIM and their absolute differences give the relationship between the image matching for different types of ground features. The multi-method image matching uses more than one image matching algorithm to check the correctness of image matching. Its use in determining optimal window size for image matching is shown. An unsquare window is used and compared with the square window. The results show that a window which has more columns than rows may improve the image matching accuracy and its speed. The conclusions are given in Chapter 9.

CHAPTER 2 INTERIOR ORIENTATION

Interior orientation is the process of registering the digital photographs into a certain pre-defined coordinate system so they can be further processed. Some of the geometric distortion on the photographs can also be removed during this process.

2.1 Definition of Coordinate Systems in Interior Orientation

There are two coordinate systems used in the interior orientation of the digital photographs. One is the pixel coordinate system which is formed by the rows and columns of the image raster matrix; and the other is the photo coordinate system which determines the real location of each pixel on the photograph. Notice that every digital photograph has its pixel coordinate defined by rows and columns if all the pixels have the same size, but only the digital photographs after interior orientation or similar process will have the photo coordinate system.

2.1.1 Pixel Coordinate System

As shown in Figure 2.1, the digital photographs are always stored in the format of a pixel matrix which is sized by its number of rows and columns. Each pixel represents a certain area. The shape of each pixel could be regular like a square or a rectangular, or irregular. Only square or rectangular pixels are used in this research. Each grid intersection point in Figure 2.1 represents the location of the center of one pixel area. The rows and columns of the pixel matrix are used to define the pixel coordinate system. According to convention, I is used to denote the X or horizontal axis from the left-hand side to the right-hand side of the image pixel matrix; J is used to denote the Y or vertical axis from top to bottom of the image pixel matrix. The directions of I, J are shown in the Figure 2.1. According to this definition shown in Figure 2.1, the pixel coordinate system is a left-hand coordinate system.

Notice that I, J are integers beginning from 0, 0 to $size\_I$-1, and $size\_J$-1 respectively. Here $size\_I$, $size\_J$ are the number of columns and rows for the directions of I and J , respectively.

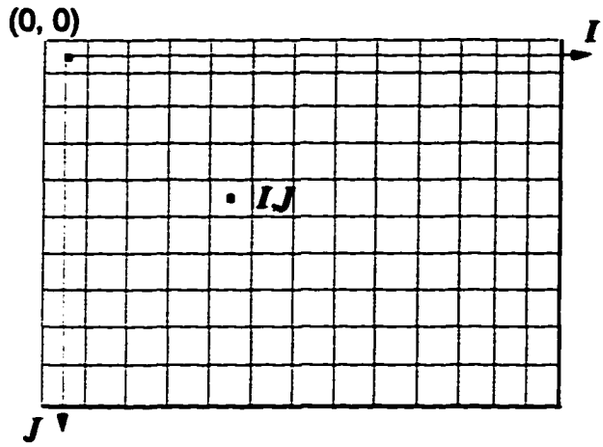


Figure 2.1 Pixel Coordinate System

The origin of the pixel coordinate system is the center of the pixel on the upper-left corner of the digital image; and that pixel is called the initial pixel.

2.1.2 Photo Coordinate System

The photo coordinate system is used to determine the location of each pixel on the digital photograph after the interior orientation. In fact the photo coordinate system is defined by the calibration data of the camera. The main objective of the interior orientation process is to register the digital photograph into the photo coordinate system. As shown in Figure 2.2, for an interiorly oriented digital photographs on the screen, x_p axis is always defined horizontally on the image from left-hand to right-hand; y_p axis is always defined vertically from bottom to top or the opposite. The definition of the photo coordinate indicates that the x_p is always displayed on the screen from the left-hand to the right-hand. The location of the origin of the photo coordinate on a photograph is determined by some pre-calibrated control points such as fiducial points in camera calibration. Assume the pixel sizes C_{xp}, C_{yp} for axes x_p and y_p respectively are of the correct sign and the photo coordinates of center of the upper-left corner pixel is x_{0p}, y_{0p}, then the relationship between pixel coordinates defined in Figure 2.1 and photo coordinates defined in Figure 2.2 on a photograph after interior orientation is:

$$x_p = C_{xp} I_p + x_{0p}$$
$$y_p = C_{yp} J_p + y_{0p} \qquad 2.1$$

After the interior orientation has been done, each pixel on a digital photograph has both pixel and photo coordinates which are related by equation 2.1. The sign of C_{yp} is determined by the type of the photo coordinate system. If the photo coordinate system is a right-hand system as shown in Figure 2.2, since the *Y* axes in these two coordinates are in opposite directions, then C_{yp} is a negative number in equation 2.1. Generally, when the photo coordinate system is a right-hand, the sign of C_{yp} is negative; when the photo coordinate system is left-hand, it is a positive number.

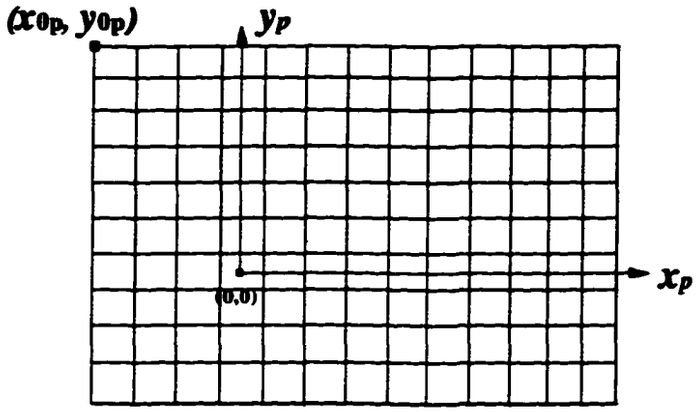


Figure 2.2 Photo Coordinate System

2.1.3 Camera Coordinate System

The camera coordinate system defines the rotation relationship between the camera and the ground coordinate system. The relationship between the camera and the photograph is fixed. If the ground control points are used in the aerial triangulation, the relationship between the photo coordinate and the ground coordinates can be directly calculated provided the photo coordinate is a right-hand system. In some situations described below, the camera coordinate system is needed:

• When it is necessary to keep the photo coordinate in the left-hand coordinate system so that it has the same orientation as the pixel coordinate of the photograph.

• When airborne GPS is used to determine the exterior orientation of the photographs, one pair of intermediate coordinates, like camera coordinates, may be needed.

The relationship between the photo coordinates and the camera coordinates can be determined by rotation and translation. In this study case the relationship between them is as simple as:

$$x_c = x_p$$
$$y_c = -y_p \qquad 2.2$$

More discussion about the camera coordinate system will be presented in the next chapter.

2.2 Acquisition of Digital Image Data

Digital images are available from many different sources. Some of the examples are satellite imagery, video imagery, scanned imagery, and so on. If the images to be processed are aerial photographs, they have to be converted to digital images first. This process is called digitization.

2.2.1 Digitizing Aerial Photographs

Digitizing aerial photographs involves sampling the photographs into pixels so that the pixel gray value of a certain area can represent the relative brightness of this area on the photograph. There are different types of scanners used in photogrammetry. Since precision is an important requirement in photogrammetry, the scanners should be able to digitize the photographs into digital images with very small pixel and very little geometric distortion. Some manufactures of the scanners used in digitizing aerial photographs are *VEXCEL* and *Wehrli*.

The scanner used in this study is an ordinary image scanner, HP *ScanJet* 3C, with a standard 75 dpi which may be not very good for image processing in soft photogrammetry. The scanned photographs are photos No. 1 and No. 2 of the campus high flight 94, as shown in Figure 2.3 and Figure 2.4. Due to the size limitation of the scanner, only five out of the eight fiducial marks are included in the scanned photographs.

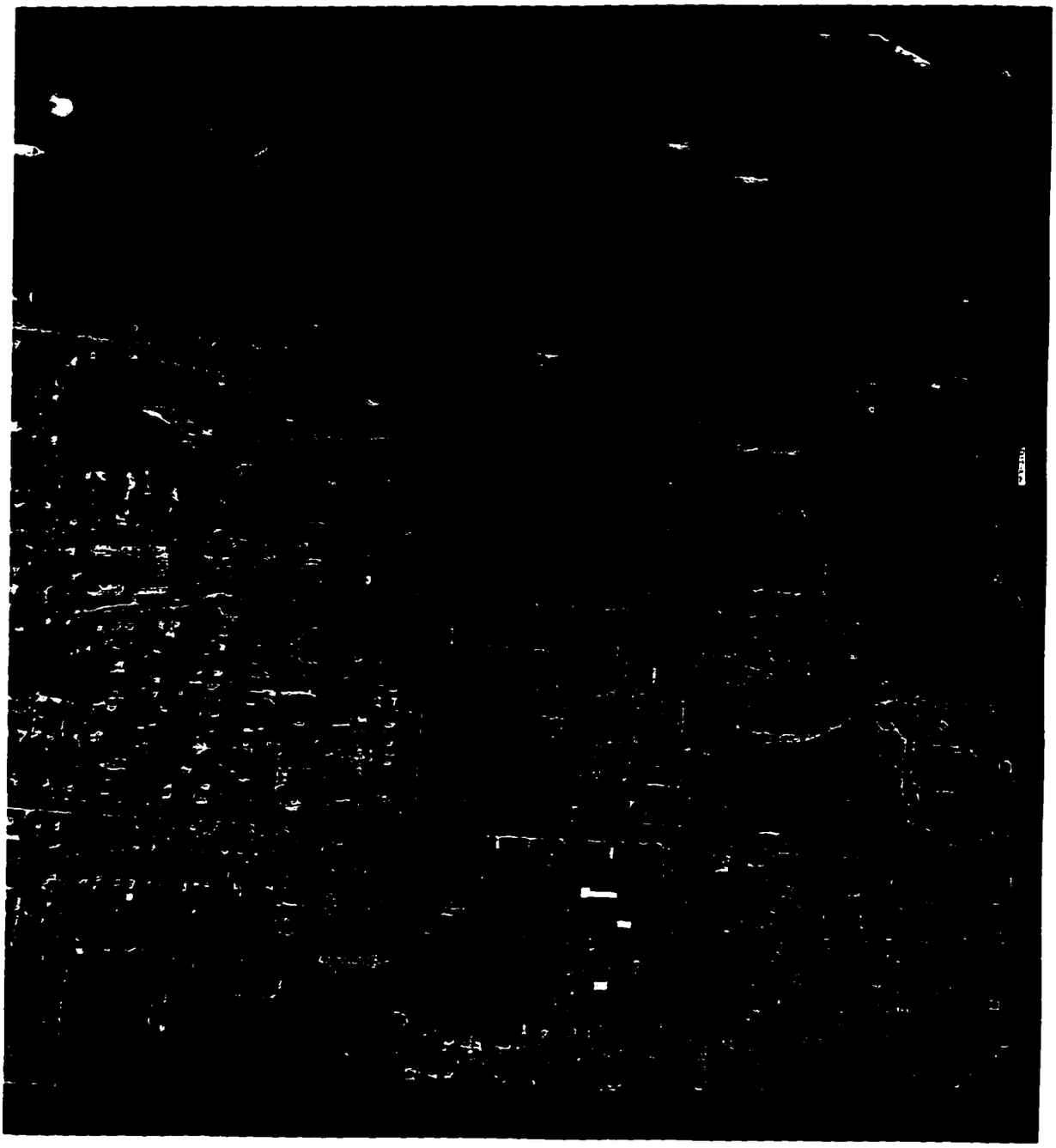


Figure 2.3 Photo No. 1 of the Campus High Flight 94

Figure 2.4 Photo No. 2 of the Campus High Flight 94

The TIFF image format is used to store the scanned images. Since the LAN format is used as the processing format, a program was written to convert all the scanned images to LAN format. The size of these two scanned images were:

For No. 1: 4361 columns by 5040 rows

For No. 2: 4380 columns by 5026 rows

2.2.2 Relationship between Pixel Size and Root Mean Square (RMS) of Interior Orientation

Because of the small pixel size, the size of output image file of the digital image is very large. If the standard size of a aerial photograph is 23 cm by 23 cm, and the pixel size is 25 μm by 25 μm, for example, the number of pixels on the scanned digital photograph would be about 84 Mb. Since the scanner HP *ScanJet 3C* is used to scan the aerial photographs used in this research, it is used to test the relationship between pixel size and the RMS in the interior orientation. The results of five different pixel sizes are shown in Figure 2.5. The RMS is from five fiducial marks on each scanned photograph.

The results shown in Figure 2.5 indicate that even when the pixel size is about 25 μm, the RMS for these five fiducial marks is around 30 μm, which means that a geometric precision less than 30 μm is not available from the photographs scanned with the HP *ScanJet 3C* scanner. But the RMS is much smaller than the nominal 75 dpi or 0.37 mm precision.

Another scanner HP *ScanJet 4C* with 600 dpi resolution was also compared. The results from the HP *ScanJet 4C* are shown in Figure 2.6. The RMS shown in Figure 2.6 is much less than that in Figure 2.5 since the higher precision scanner HP *ScanJet 4C* was used. The lowest RMS here is about 10 μm at the pixel size of about 40 μm. An image with the pixel size of about 28 μm was tried but failed because of the software SoftPlotter may not be able to handle some a big file in the interior process. But according to the trend in Figure 2.6, a lower RMS than 10 μm is available by using the HP *ScanJet 4C* scanner. The conclusion here is that the higher resolution or smaller pixel size than the nominal value can be obtained from a low precision scanner, but the geometric accuracy still remains low, which means the distortion of the scanned image is determined by the geometric precision of the scanner.

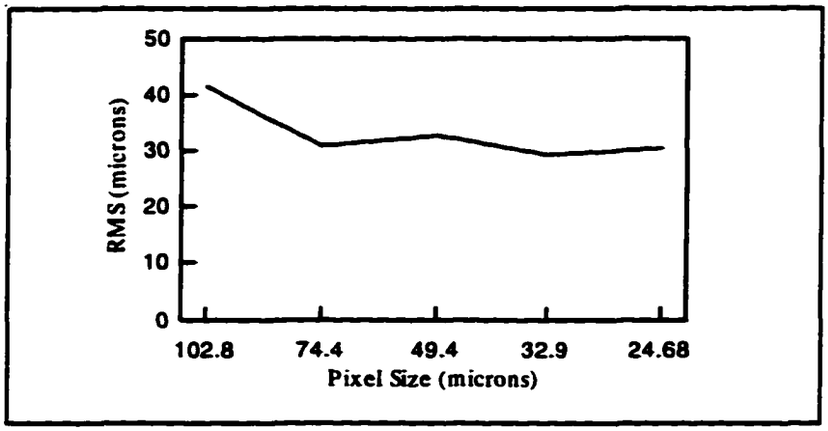


Figure 2.5 RMS vs. Pixel Size for HP *ScanJet 3C*

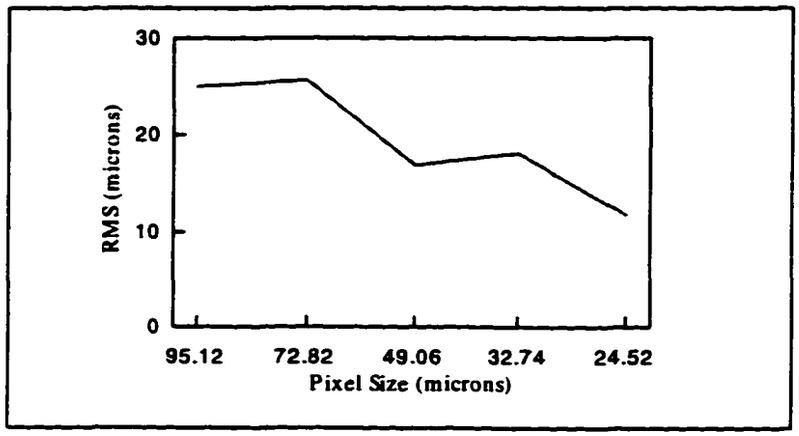


Figure 2.6 RMS vs. Pixel Size for HP *ScanJet 4C*

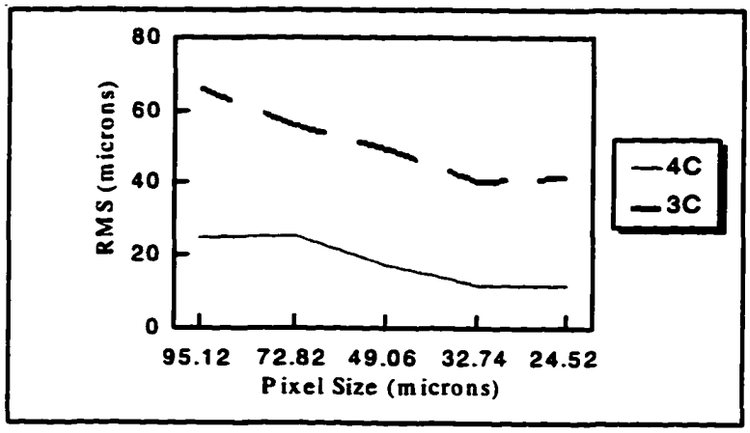


Figure 2.7 HP *ScanJet 4C* vs. HP *ScanJet 3C*

Figure 2.7 shows the relationship between the RMS and the different pixel sizes for both 3C and 4C.

2.2.3 Content of the Header of the Modified LAN Images

The LAN image format is used in this research to store the images involved and all the information about the images and the stereo model (if it exists). All the information about the LAN images are stored in the header which is the first 128 bits in the image data. In addition to the normal LAN header, some new items are added to the header to provide more information about the images. Table 2.1 shows the content of the modified LAN header.

Table 2.1 Content of the Modified LAN Header

NAME	WIDTH (BYTES)	TYPE	REMARKS
HDWORD	6	CHAR	Head word, always "HEAD74"
UNUSED1	6	CHAR	Unused bytes
IPACK	2	SHORT INT	Number of bits per sample
NBANDS	2	SHORT INT	Number of bands
ICOLS	4	LONG	Number of columns
IROWS	4	LONG	Number of rows
XSTART	4	LONG	X start position of this image in the map database
YSTART	4	LONG	Y start position of this image in the map database
UNUSED2	15	CHAR	Unused bytes
COMPRESSION	1	CHAR	Compression type
DIR	2	CHAR	Flight and reference directions
SCALE	2	SHORT INT	Ratio between pixel gray value and real height on DEM
MAXIMUM	4	LONG	Maximum height on DEM * SCALE
RATIO	8	DOUBLE	Ratio used in the scale equalizing
XL	8	DOUBLE	X ground coordinate of the camera
YL	8	DOUBLE	Y ground coordinate of the camera
ZL	4	FLOAT	Z ground coordinate of the camera
FOCAL	4	FLOAT	Principal Distance
MAPTYPE	2	SHORT INT	Type of map unit
NCLASS	2	SHORT INT	Number of class
MINIMUM	4	LONG	Minimum height on DEM * SCALE
ANGLE	8	DOUBLE	Angle rotated in epipolar rectification
MARK	2	SHORT INT	Type of photograph in the image matching
IAUTYP	2	SHORT INT	Unit of area
ACRE	4	FLOAT	Area of this image
XMAP	4	FLOAT	X initial location of this image
YMAP	4	FLOAT	Y initial location of this image
XCELL	4	FLOAT	Pixel size in X direction
YCELL	4	FLOAT	Pixel size in Y direction

The standard LAN image format is used for *Landsat* images. In the process of soft photogrammetry, large quantities of information about the status of the process and many parameters, such as the exterior orientation parameters and focal length of the camera, are needed. In SoftPlotter and Socet Set, the types of information are kept in separate project files. In this research, all the information needed is kept in the modified LAN header. There are some bits in the header that are not used. The unused space in the LAN header can be used to store the information for the soft photogrammetric process. More bits can be added to the end of the image data if the unused space in the header is not enough, just like the TIFF image format. The advantage of this is that image files have almost all of the information about the processing. Also, the images with the modified LAN header can still be recognized by software packages like IMAGINE, they can work just like the standard LAN format images.

Values of some of the items in Table 2.1 are:

IPACK:

- 0 for 8 bits per sample
- 2 for 16 bits per sample
- 4 for 32 bits per sample

MARK:

- 0 for the raw photographs
- 1 for the photographs after interior orientation
- 2 for rectified photographs
- 3 for photographs in a stereo pair with the same scale
- 4 for DEM before post-processing
- 5 for DEM represented by the minimum and the difference between the minimum and the maximum of the pixels in the image space
- 6 for DEM in the ground space
- 7 for orthophotographs in the ground coordinate system
- 8 for relatively oriented photographs in a stereo pair

- 9 for a pair of photographs in epipolar coordinates system

2.2.4 Initialization of the Header of the LAN Images

Before doing the interior orientation, the header of the LAN images needs to be initialized for the items that are not determined by the images. Table 2.2 shows the initialization values.

The unit of the coordinates in the image space is millimeters. For location of the photograph in the ground coordinate system, the unit is meters.

Table 2.2 Initialization of the Modified LAN Header

NAME	REMARKS
HDWORD	Head word, always "HEAD74"
UNUSED1	Unused bytes
IPACK	Determined by image
NBANDS	Determined by image
ICOLS	Determined by image
IROWS	Determined by image
XSTART	Set to 0
YSTART	Set to 0
UNUSED2	Unused bytes
COMPRESSION	Set to 0
DIR	Set to 'EE'
SCALE	Set to 1
MAXIMUM	Set to 1
RATIO	Set to 1
XL	No initialization
YL	No initialization
ZL	No initialization
FOCAL	No initialization
MAPTYPE	Set to 0
NCLASS	No initialization
MINIMUM	Set to 1
ANGLE	Set to 0
MARK	Set to 0
IAUTYP	No initialization
ACRE	No initialization
XMAP	No initialization
YMAP	No initialization
XCELL	No initialization
YCELL	No initialization

2.3 Interior Orientation of the Digital Aerial Photographs

The location of a point on the digital photograph depends on how the scanning is done, but its photo coordinate is independent of the scanning process. The goal of interior orientation of the digital photographs is to (1) get the photo coordinates of all the pixels on the digital photograph (2) remove some of the distortion caused by scanning and the camera imaging system.

2.3.1 Determination of Transformation Parameters in the Interior Orientation

The transformation between the digital photographs before and after the interior orientation can be linear or non-linear according to the number of control points and the precision requirement throughout the photograph. In this research 6 parameter Affine transformation shown in equation 2.3 is used, which is sufficient for many applications. The Affine transformation equation is:

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} I \\ J \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \qquad 2.3$$

Here a, b, c, d, x_0, y_0 are 6 transformation parameters; I, J are the pixel coordinates before the interior orientation. Each fiducial point or known control point can give two equations, so at least 3 known points are needed to find the 6 parameters. More than three points are recommended so that the least square method can be used to solve the equation and check to ensure that the measurements are correct.

To get these 6 unknown parameters from m known control points, the following equations are established:

$$\begin{bmatrix} x_1 \\ y_1 \\ \cdots \\ x_m \\ y_m \end{bmatrix} = \begin{bmatrix} I_1 & J_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_1 & J_1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ I_m & J_m & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_m & J_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ x_0 \\ c \\ d \\ y_0 \end{bmatrix} \qquad 2.4$$

Solve the equation 2.4 for these 6 unknown parameters. The standard errors can be used to check the quality of the measurements. In this process, some of the geometric distortion caused by scanning and the camera imaging system can be removed.

The camera radial distortion should also be removed when observing the control and tie points. All the observations should be corrected before applied.

2.3.2 Transformation Equation for the Interior Orientation

To perform the interior orientation, the pixel coordinate system must be used. The pixel coordinates on the photograph after interior orientation are given by equation 2.3 and the equation below

$$x_p = C_{xp} I_p + x_{0p}$$
$$y_p = C_{yp} J_p + y_{0p}$$

Where C_{xp}, C_{yp} are the pixel sizes in x_p and y_p directions respectively; I_p, J_p are the pixel coordinates on the interiorly oriented photographs; and x_{0p}, y_{0p}, are the initial photo coordinates of the interiorly oriented photographs in the photo coordinate. The equation 2.3 can be written in terms of pixel coordinates as:

$$\begin{bmatrix} I_p \\ J_p \end{bmatrix} = \begin{bmatrix} \frac{a}{C_{xp}} & \frac{b}{C_{xp}} \\ \frac{c}{C_{yp}} & \frac{d}{C_{yp}} \end{bmatrix} \begin{bmatrix} I \\ J \end{bmatrix} + \begin{bmatrix} \frac{x_0 - x_{0p}}{C_{xp}} \\ \frac{y_0 - y_{0p}}{C_{yp}} \end{bmatrix} \qquad 2.5$$

$$I = 0, \cdots, size\_I - 1; J = 0, \cdots, size\_J - 1$$

Then the location (I, J) of each pixel on the raw photo can be transformed to the photo coordinate (I_p, J_p). Before performing the interior orientation, some parameters in equation 2.5 need to be determined.

2.3.3 Determination of Pixel Sizes

From these 6 parameters the pixel size on the digital images after interior orientation are calculated as:

$$C_{xp} = \sqrt{a^2 + b^2}$$
$$C_{yp} = \frac{ad - bc}{|ad - bc|} \sqrt{c^2 + d^2} \qquad 2.6$$

The sign of C_{yp} is determined by the product of parameters a and d minus the product of b and c. If the result is positive, the photo coordinate is of the same coordinate orientation as the pixel coordinate, that is the left-hand coordinate system, then C_{yp} is a positive number. Otherwise the photo coordinate system is a right-hand coordinate system like the one shown in Figure 2.2, and C_{yp} is negative.

2.3.4 Determination of Initial Position in Photo Coordinate

The initial photo coordinates x_{0p}, y_{0p} in equation 2.5 need to be determined before the interior orientation can be performed. Since x_{0p}, y_{0p} are the photo coordinates of the center of the upper-left corner pixel on the photograph after interior orientation and the x_p is defined as shown in Figure 2.2, only the minimum of x_p and minimum or maximum of y_p on the interiorly oriented photograph can be taken as the values of x_{0p}, y_{0p}. Therefore, from equation 2.5, the limits of the interiorly oriented photograph are:

$$\begin{aligned} x_{\min} &= \min(aI + bJ + x_0) \\ y_{\min} &= \min(cI + dJ + y_0) \\ x_{\max} &= \max(aI + bJ + x_0) \\ y_{miax} &= \max(cI + dJ + y_0) \\ I &= 0, size\_I - 1; \quad J = 0, size\_J - 1 \end{aligned} \tag{2.7}$$

Where ***size_I***, ***size_J*** are the number of columns and rows on the raw photograph, corresponding to the size of the photograph before the interior orientation; $x_{\min}$, $x_{\max}$, $y_{\min}$, $y_{\max}$ are the limits of the photograph after interior orientation. Then the initial photo coordinates of the interiorly oriented photograph is given by:

$$\begin{aligned} x_{0p} &= x_{\min} \\ y_{0p} &= y_{\max} \quad if \quad C_{yp} < 0 \\ y_{0p} &= y_{\min} \quad if \quad C_{yp} > 0 \end{aligned} \tag{2.8}$$

2.3.5 Determination of Photograph Size after Interior Orientation

From equation 2.7 the pixel size of the interiorly oriented photograph is given as:

$$\begin{aligned} size\_x_p &= \frac{x_{\max} - x_{\min}}{C_{xp}} \\ size\_y_p &= \frac{y_{\max} - y_{\min}}{|C_{yp}|} \end{aligned} \tag{2.9}$$

Where $size\_x_p$, $size\_y_p$ are the number of columns and rows on the interiorly oriented photographs. Therefore by using equation 2.5 through 2.9 for each pixel on the raw digital photograph, the corresponding interiorly oriented pixel and photo coordinates can be calculated. Since the interiorly oriented photographs are also in the pixel coordinate, and the pixel coordinates from the equation 2.5 may not be integers, resampling processing is needed in the interior orientation.

2.3.6 Resampling Equation in the Interior Orientation

Resampling is a interpolation process used to find the pixel gray values other than the center of the pixel on the digital images. Different resampling methods will be introduced in the next section. In this section only the resampling equation is given. Resampling equation is the reverse equation of the interior orientation transformation equation shown in equation 2.5. For a given pixel on the interiorly oriented photograph, the corresponding pixel location on the raw photograph can be found. If that pixel location is not an integer, then the pixel gray value on that location needs to be interpolated from the surrounding pixels. The resampling equation for interior orientation is

$$\begin{bmatrix} I \\ J \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} \left(\begin{bmatrix} C_{xp} I_p \\ C_{yp} J_p \end{bmatrix} + \begin{bmatrix} x_{0p} - x_0 \\ y_{0p} - y_0 \end{bmatrix} \right) \qquad 2.10$$

$$I_p = 0, \cdots, size\_x_p - 1; J_p = 0, \cdots, size\_y_p - 1$$

For each pixel (I_p, J_p) on the photograph after interior orientation, its corresponding pixel location (I, J) on the raw photograph can be computed with equation 2.10; and the pixel gray value is available by means of interpolation.

2.4 Resampling in Digital Image Transformation

Resampling is the process used to obtain the pixel gray value on the imagery after transformation from the original imagery by means of interpolation. Resampling is needed in almost all the digital image transformation. The resampling in the interior orientation process is one of the examples. Assume the pixel gray value for the raw photograph is $g(I, J)$ for the pixel on (I, J); and the pixel gray value for the photograph after interior orientation is $g_p(I_p$,

J_p). The resampling equation is the same as the equation 2.10. Since the pixel position on the raw photograph from equation 2.10 may not be integers, equation 2.10 is rewritten as:

$$\begin{bmatrix} I' \\ J' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} \left(\begin{bmatrix} C_{xp} I_p \\ C_{yp} J_p \end{bmatrix} + \begin{bmatrix} x_{0p} - x_0 \\ y_{0p} - y_0 \end{bmatrix} \right) \qquad 2.11$$

$$I_p = 0, \cdots, size\_x_p - 1; J_p = 0, \cdots, size\_y_p - 1$$

Where I', J' are not integers. Once I', J' are calculated, $g_p(I_p, J_p)$ can be determined by interpolation of the neighbor pixels around location (I', J') on the raw photograph. There are many different interpolation methods used in the resampling process. In this study, only nearest neighbor, bilinear interpolation, and distance weighted interpolation methods are used.

2.4.1 Nearest Neighbor Interpolation

Nearest neighbor is the simplest and fastest resampling method to get the pixel gray value on the photograph after the transformation. Its idea is to take the value of the nearest pixel as the interpolated value. So actually there is no interpolation in the nearest neighbor resampling. As shown in Figure 2.8 the distances between (I', J') of location A and the other four surrounding pixels are calculated and compared. The pixel gray value with the shortest distance is taken as the value of pixel (I', J'). According to equations 2.5, 2.11, assume the shortest distance is between the points A and F shown in Figure 2.8, then the pixel gray value on the interiorly oriented photograph is given by

$$g_p(I_p, J_p) = g(I', J') = g(I, J+1) \qquad 2.12$$

So the pixel gray value of the nearest pixel F is taken as the pixel gray value on the transformed photograph. This algorithm has the least computation, so it is the fastest one. It can give satisfactory results form some applications, especially for those which need to keep the pixel gray values.

2.4.2 Bilinear Interpolation

In the bilinear interpolation the four surrounding pixels are involved. As shown in Figure 2.8, the pixel gray value on location (I', J') on the raw image can be determined by the following bilinear interpolation equation:

$$\begin{aligned} g(I', J') = {} & (g(I, J)(I + 1 - I') + g(I + 1, J)(I' - I))(J + 1 - J') + \\ & (g(I + 1, J + 1)(I' - I) + g(I, J + 1)(I + 1 - I'))(J' - J) \end{aligned} \qquad 2.13$$

and the pixel gray value for the point (I_p, J_p) on the transformed photograph is:

$$g_p(I_p, J_p) = g(I', J')$$

since they are the same point on the two photographs before and after transformation.

2.4.3 Distance Weighted Interpolation

The inverses of the distances between the location to be interpolated and the surrounding pixels are used as weights in the distance weighted interpolation. Two types of interpolation are used in the distance weighted interpolation. They are bilinear and cubic convolution. The weight used in bilinear and cubic convolution interpolation for each pixel involved is inversely proportional to the distance between the location to be interpolated and the surrounding pixels.

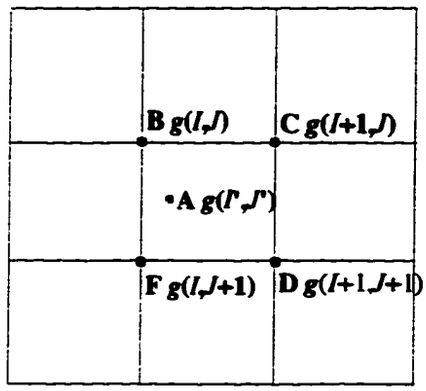


Figure 2.8 Inpterpolation in Resampling

As shown in Figure 2.8, assume for given pixel (I_p, J_p) on the transformed image, (I', J') are the pixel coordinates calculated from equation 2.11. If the location of (I', J') is similar to that shown in Figure 2.8, the algorithms used in the two types of interpolation are bilinear and cubic convolution interpolation.

2.4.3.1 Bilinear Interpolation

In this case the four surrounding pixels are used in the interpolation, just like that one discussed above. The inverse distances are used as the weights in the interpolation. According to Figure 2.8, numbering the distances between the point A and others as D_i, i = 1,2,3,4, then the pixel gray value of the point A is determined by

$$g_p(I_p, J_p) = g(I', J') = \frac{\frac{g(I,J)}{D_i^k} + \frac{g(I+1,J)}{D_i^k} + \frac{g(I+1,J+1)}{D_i^k} + \frac{g(I,J+1)}{D_i^k}}{\sum_{i=1}^{4}(1/D_i^k)} \qquad 2.14$$

Where the k is a constant and normally its value is larger than 2. 2 is used her for the k. According to the equation 2.14 it is known that the shorter the distance, the larger the weight of that pixel has in the interpolation. The weight will be infinite for one pixel if the distance to that pixel is 0. In the case of zero distance, the gray value of that pixel should be used, similar to the nearest neighbor interpolation.

2.4.3.2 Cubic Convolution Interpolation

The only difference between the bilinear and the cubic convolution is that more pixels are used in the cubic convolution interpolation. In this case, there are 16 surrounding pixels involved in the interpolation. Number all the pixels and their distances between the point in the Figure 2.8 as g_i, D_i; i = 1, 2, ..., 16. Then the pixel gray value of the point A is calculated by

$$g_p(I_p, J_p) = \frac{\sum_{i=1}^{16} \frac{g_i}{D_i^k}}{\sum_{i=1}^{16}(1/D_i^k)} \qquad 2.15$$

Where k is the same constant as that in the bilinear interpolation.

2.5 Results of the Interior Orientation

The interiorly oriented images of Figure 2.3 and Figure 2.4 are given in this section. The results include the six transformation parameters, the two interiorly oriented photographs, and the subset photographs from the interiorly oriented photographs in the overlapping area.

2.5.1 The Six Transformation Parameters

Before determining the 6 unknown transformation parameters, the location of the five fiducial marks on the scanned photographs must be measured. Table 2.3 shows the calibration data and the measurements of the fiducial points.

Substituting the values from Table 2.3 in equation 2.4 and solving the equation to get the six Affine transformation parameters as shown in Table 2.4.

The pixel sizes in x_p and y_p direction after interior orientation can calculated from the Table 2.4 with the parameters a, b and c, d. The pixel size will be shown in the section 2.5.3.

Table 2.3 Measurements of the Fiducial Marks and the Calibration Data

Mark No.	Calibration Data (mm)		Measurements for Photo No. 1		Measurements for Photo No. 2	
	X	Y	Pixel X	Pixel Y	Pixel X	Pixel Y
1	109.989	109.994	4185.06	203.56		
2	110.001	-109.999	4194.94	4646.44		
3	111.995	-0.001	4230.56	2424.06		
4	-0.008	111.998	1966.06	166.06	2291.81	175.06
5	0.004	-112.003	1976.94	4688.94	2291.81	4724.06
6	-109.994	-109.999			62.69	4681.19
7	-110.007	109.996			63.69	212.94
8	-112.002	0.002			24.31	2447.19

Table 2.4 Values of the 6 Parameters for both Photographs

Name of the Parameter	Parameters for No. 1	Parameters for No. 2
a	0.04 957986	0.04936983
b	-0.00011216	0.00000820
x_0	-97.47634405	-113.16788498
c	-0.00006239	0.00005287
d	-0.04952097	-0.04923854
y_0	120.33290581	120.48901416

2.5.2 Interior Orientation Results

The photographs after interior orientation are shown in Figures 2.9 and 2.10. It does not seem to be much difference between the raw photographs shown in Figures 2.3 and 2.4 and the photographs after interior orientation shown in Figures 2.9 and 2.10. But they are in totally different systems. In the interiorly oriented photographs, the photo coordinates of any pixel are available from its pixel coordinates; and the pixel sizes in x and y directions are also available. The initialization of the LAN header of each photograph is also done after the

interior orientation. So the photographs after interior orientation are in the coordinate system that is connected to the camera.

2.5.3 Subsets of the Interiorly Oriented Photographs

Since only the overlapping area can be used in the stereo model, the selected subsets in the common area for both photographs are used.

In the process of getting the subset of a photograph, the photo coordinates of one pixel on the subset photographs should remain the same coordinates as they are on the original photograph. The formula used to obtain subset from a given photograph is:

$$\begin{aligned} I'_p &= I_p + I_{01} \\ J'_p &= J_p + J_{01} \\ x'_{0p} &= x_{0p} + C_{xp} I_{01} \\ y'_{0p} &= y_{0p} + C_{yp} J_{01} \\ size\_x'_p &= I_{02} - I_{01} \\ size\_y'_p &= J_{02} - J_{01} \end{aligned} \qquad 2.16$$

Where I'_p, J'_p are the pixel coordinates on the subset image; I_p, J_p are the pixel coordinates on the original image; I_{01}, J_{01} and I_{02}, J_{02} are the coordinates of the upper-left and lower-right pixels of the subset images on the original image respectively; x_{0p}, y_{0p} are the initial position of the subset image; $size\_x'_p$, $size\_y'_p$ are the new size. Since all the pixel coordinates in equation 2.16 are integers, there is no need for resampling.

The subset photographs are shown in Figures 2.11 and 2.12.

The information about the header of these two subset photographs are shown in Table 2.5.

Table 2.5 Header Information about the Subset Photographs

Name of the Parameter	Parameters for No. 1	Parameters for No. 2
x_{0p}	-13.755544	-106.453590
y_{0p}	105.476608	105.948967
$size\_x_p$	2600	2715
$size\_y_p$	3700	3700
Pixel size in x	0.049580	0.049370
Pixel Size in y	-0.049521	-0.049239

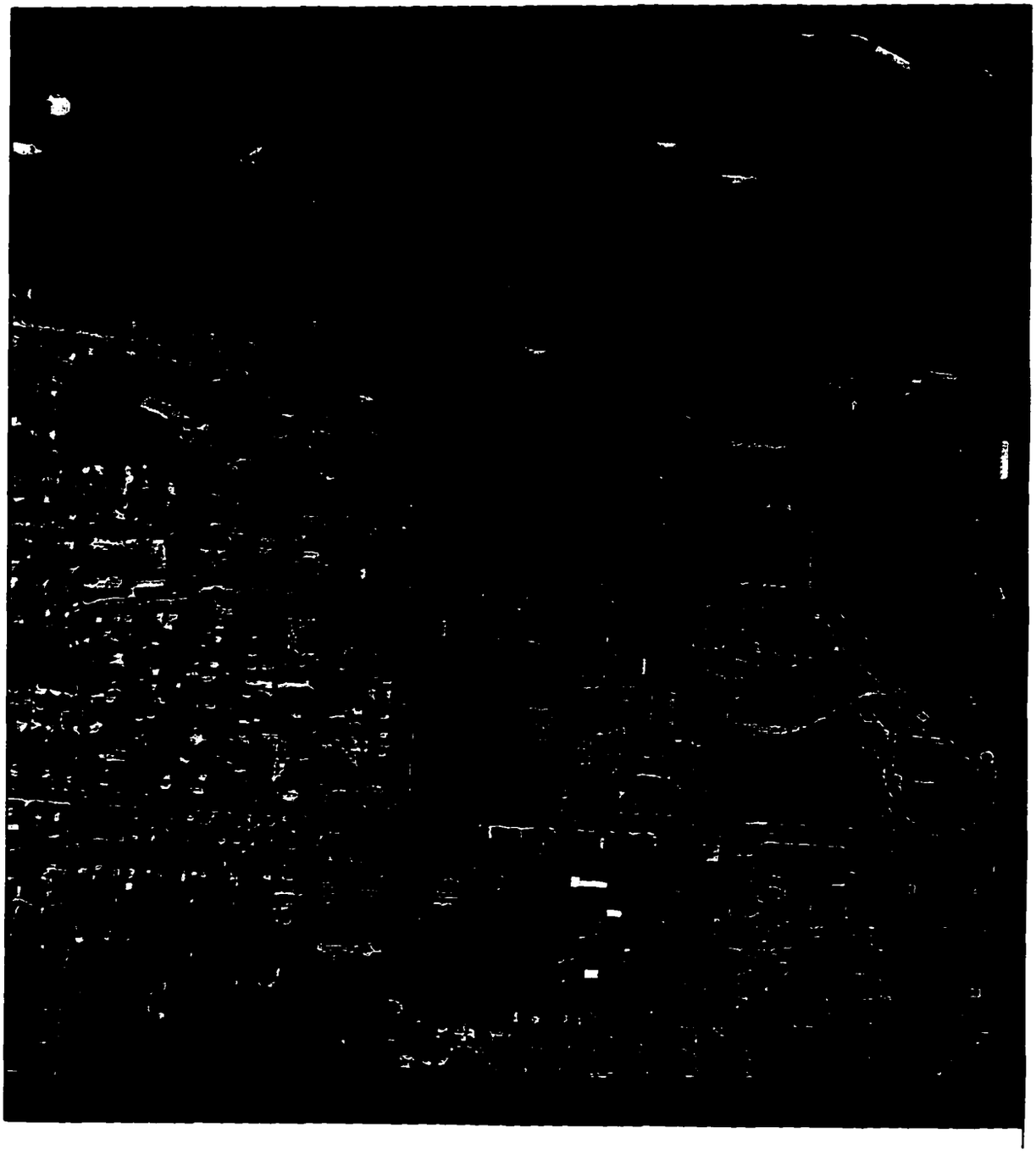


Figure 2.9 Photo No. 1 after Interior Orientation

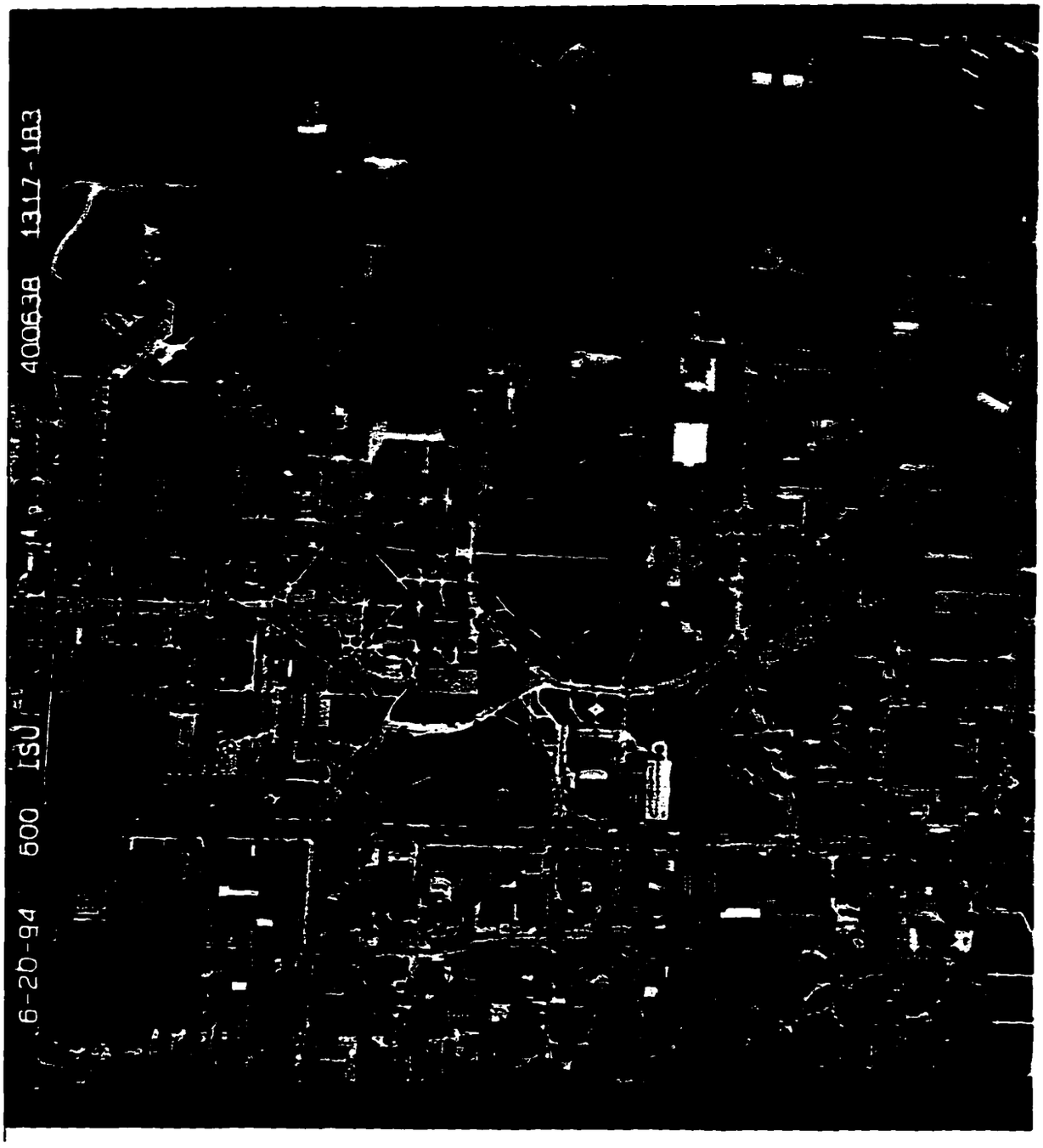


Figure 2.10 Photo No. 2 after Interior Orientation

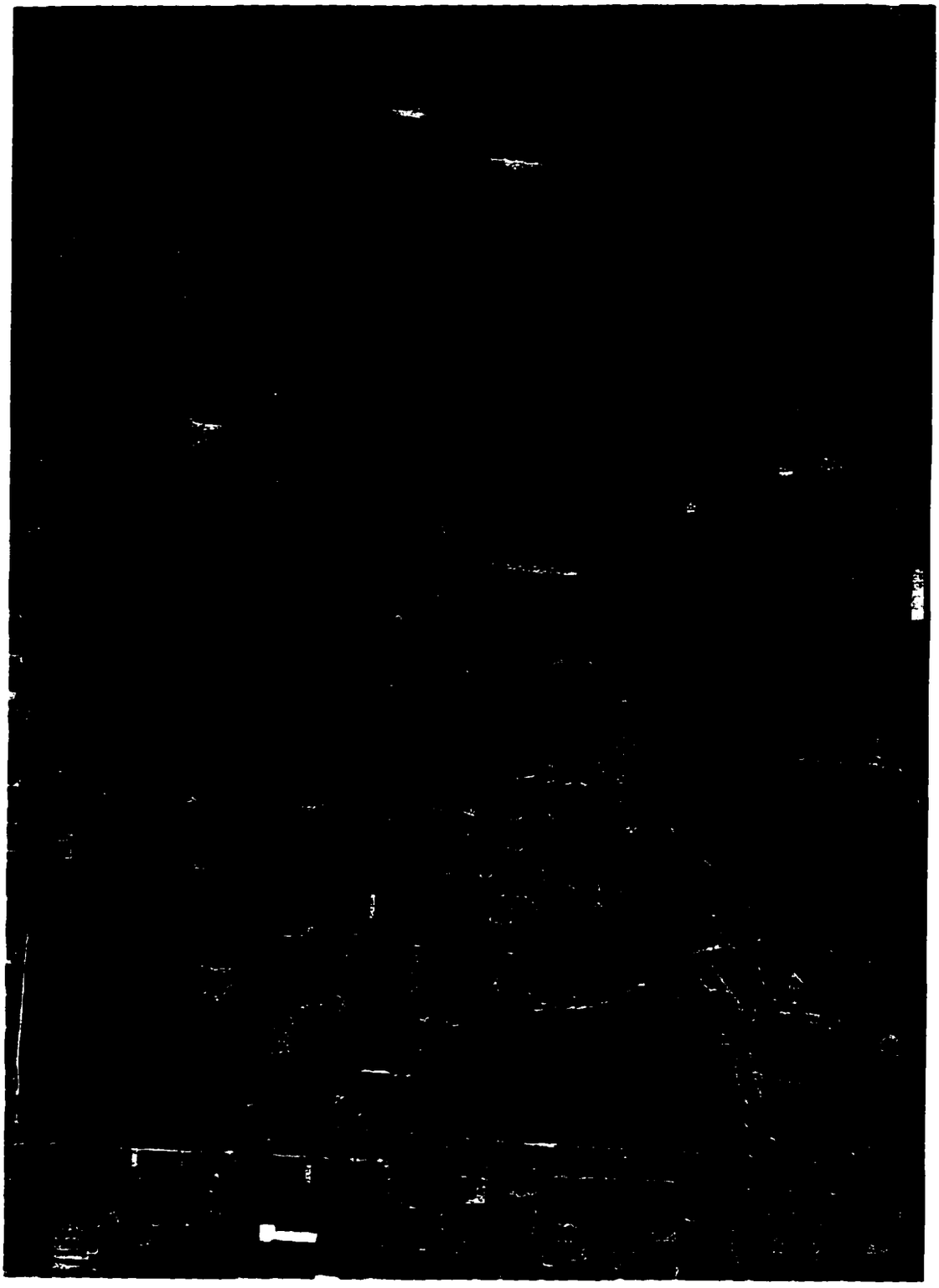


Figure 2.11 Overlapping Area from Photo No. 1

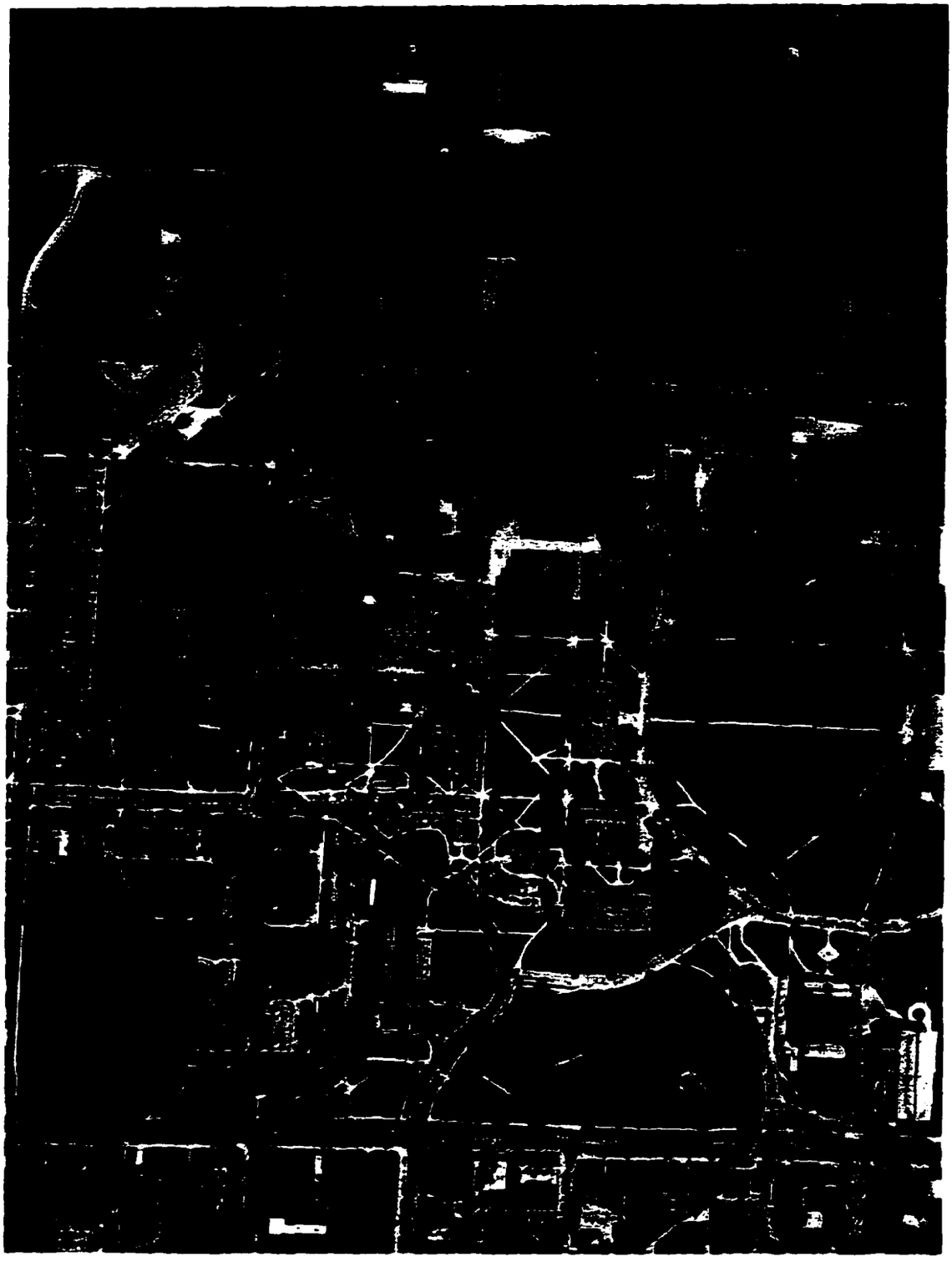


Figure 2.12 Overlapping Area from Photo No. 2

2.6 Conclusions

The details of the process of interior orientation and some related topics have been discussed above. They include the modified LAN image format, scanning process, resampling process in soft photogrammetry, and interior orientation.

The modified LAN image format was presented and will be used in this research for storing the information about all the kinds of images, such as raw, rectified, DEM and orthophotos. All the information about the digital image will be kept in the modified header of the LAN file, which makes the management of the digital images easy.

As shown in Figure 2.7, the more accurate the scanner, the less the RMS for the digital images scanned at the same file size. Due to the limitation of pixel size, to a certain extent more accurate scanner will not produce higher geometrically accurate digital image for a certain pixel size. The other limitation for the ordinary scanner is that the size of the scanner may not be large enough to accommodate the entire aerial photograph.

The complete interior orientation process was presented by using the example of two scanned aerial photographs, and two subsets of the interiorly oriented photographs are obtained for the same overlapping area.

The programs written can be applied to accomplish each step in the interior orientation discussed in this chapter.

CHAPTER 3 SINGLE PHOTOGRAPH RECTIFICATION

Single photograph rectification is the process used to correct errors on each aerial photograph caused by the tilt of the camera when they were taken, and brings the photo coordinate system parallel to the ground coordinate system. Since the aerial photographs may not be taken exactly vertically, the rectification process is necessary before image matching. In rectifying the aerial photographs, it is brought back to vertical position and the distortion caused by the tilt of the camera is removed. For each aerial photograph, there are six parameters used to determine the exterior orientation of the photograph. They are three angles ω, ϕ, κ and ground coordinates X, Y, Z of the camera nodal point which will be discussed in the next section. The single photograph rectification process is also called exterior orientation. Similar to the interior orientation, the exterior orientation process connect the aerial photographs to the ground coordinate system.

Unlike the stereo rectification discussed in the next chapter, the single photograph rectification for each photograph is separately performed, or it is only applied to single photograph.

In order to perform the single photograph rectification, some triangulation results from different software packages are used and compared.

3.1 Definition of Coordinate Systems and Angles

There are two coordinate systems involved in the process of rectification, they are camera coordinate system and ground coordinate system. The relationship between the directions of these two coordinate systems is defined by three independent angles.

3.1.1 Camera Coordinate System

Camera coordinate system has the same origin as the photo coordinate system. As shown in Figure 3.1, it is a right-hand coordinate. Unlike the photo coordinate system, the camera coordinate system is a three dimensional coordinate system, which is connected to the orientation of the camera in or attached to the airplane. Assuming that the x axis of the camera coordinate system has the same direction as the photo coordinate system. Since the

camera coordinate system is a three dimensional coordinate system, the z coordinate of the origin is $-f$. Where f is the principal distance of the camera.

In the simplest case, the photo coordinates can be converted to camera coordinates with the following equation:

$$\begin{aligned} x_c &= x_p \\ y_c &= y_p \ when \quad C_y < 0 \quad or \\ y_c &= -y_p \ when \quad C_y > 0 \\ z_c &= -f \end{aligned} \qquad 3.1$$

Where x_p, y_p are photo coordinates and x_c, y_c, z_c are camera coordinates.

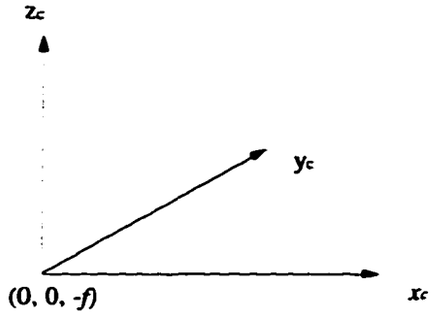


Figure 3.1 Camera Coordinate System

3.1.2 Ground Coordinate System

The ground coordinate system used in map making can be in either UTM, SPC or any local system. Generally speaking, the ground coordinate system is a right-hand coordinate system and defined in Figure 3.2.

Conventionally, the X axis is in the direction of east and the Y axis in the direction of north; the Z axis is in vertical direction.

3.1.3 Rectified Coordinate System

The rectified coordinate system (x_r, y_r, z_r) has the same orientation as the ground coordinate system, the only difference is that the origin of rectified coordinate system is the

same as that of the camera coordinate system. Unlike the ground coordinate system, the rectified coordinate system is more related to the image space of the camera.

The value of z_r/f defines the ratio of the scale of the rectified photograph to the scale of the photograph before rectification.

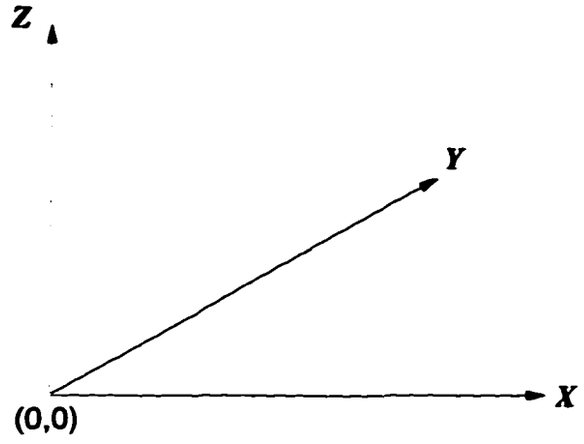


Figure 3.2 Ground Coordinate System

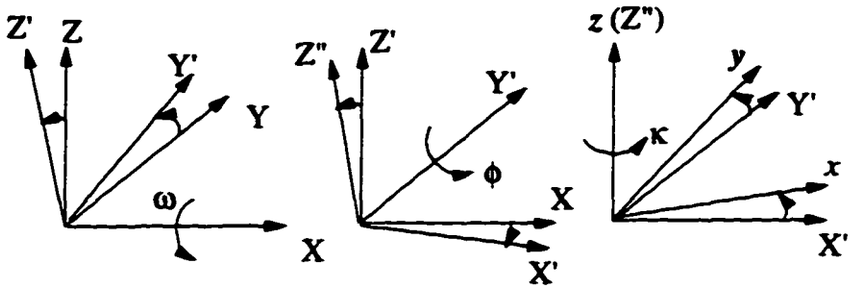
Usually, $z_r = -f$ in the rectification for same scale rectification. Therefore, the relation between rectified and ground coordinate systems is

$$\begin{aligned} x_r &= X - X_0 \\ y_r &= Y - Y_0 \\ z_r &= Z - Z_0 \end{aligned} \qquad 3.2$$

Where X_0, Y_0, Z_0 are the ground coordinates of the camera nodal point.

3.1.4 Definition of the Angles of ω, ϕ and κ

The three angles are defined as the (a), (b) and (c) in Figure 3.3. The first rotation is the angle ω around the X axis of the ground coordinate to obtain the new coordinate $XY'Z'$; then the new coordinate rotated around axis Y' by the angle ϕ becomes the next coordinate $X'Y''z$; the last rotation is around z by the angle of κ; then the coordinate is turned into camera coordinates from the ground coordinates. Notice that these angles are right-hand rotation angles and the rotation order can not be changed.



(a) Rotation around X axis

(b) Rotation around Y' axis

(c) Rotation around z (Z") axis

Figure 3.3 Definition and order of ω, ϕ and κ

Assuming R_ω, R_ϕ, R_κ are the transformation matrices for the above rotation. In terms of ω, ϕ, κ, these matrices are determined by equation 3.3. Then the rotation matrix from the ground coordinates to the camera coordinates is computed by equation 3.4.

The directional relationship between camera coordinates (x_c, y_c, z_c) and the ground coordinates (X, Y, Z) is determined by equation 3.5 for all the pixels on the photograph. Where $\boldsymbol{R}$ is a unit orthogonal matrix from equation 3.4; (X_0, Y_0, Z_0) are the coordinates of the camera nodal point; s is the scale factor. The value of s is different from pixel to pixel on the photograph, which is the function of ground height.

Notice that the order of these three rotation can not be changed, especially in the case of large rotation angles; otherwise the value of the matrix $\boldsymbol{R}$ could be much different. In the case of small angles less than 3°, there might not be much difference. The order of the rotation in this research is defined by Figure 3.3. In fact, what is really needed here is the transformation matrix $\boldsymbol{R}$. If the $\boldsymbol{R}$ in equation 3.5 can be found, then the relationship between the image and the ground coordinate is available, and it is sufficient to do the single photograph rectification.

$$R_\omega = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix}$$

$$R_\phi = \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix} \qquad 3.3$$

$$R_\kappa = \begin{bmatrix} \cos\kappa & \sin\kappa & 0 \\ -\sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = R_\kappa R_\phi R_\omega = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \qquad 3.4$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} s \qquad 3.5$$

3.2 Aerotriangulation

Aerotriangulation is the process used to determine the six exterior parameters for each aerial photograph in strip or block of photographs and the position of the pass points or other unknown points on the photographs.

3.2.1 Determination of the Exterior Orientation Parameters

The exterior orientation parameters of each photograph are X_0, Y_0, Z_0, ω, ϕ, κ which can be determined by using ground control points with the co-linearity condition. The coplanerity condition can be used as constraints for checking the relative orientation at the same time. For given ground control points X_i, Y_i, Z_i and its image point on photo x_i, y_i, I = 1, ..., n the following co-linearity equation is used.

In equation 3.6, f is the principal distance; r_{11}, ..., r_{33} are the elements from the rotation matrix R in equation 3.4 which is the function of ω, ϕ and κ. Since each control point can give two equations, at least three control points are needed to solve the six unknown parameters.

$$x_i = \frac{r_{11}(X_i - X_0) + r_{12}(Y_i - Y_0) + r_{13}(Z_i - Z_0)}{r_{31}(X_i - X_0) + r_{32}(Y_i - Y_0) + r_{33}(Z_i - Z_0)}(-f)$$
$$y_i = \frac{r_{21}(X_i - X_0) + r_{22}(Y_i - Y_0) + r_{23}(Z_i - Z_0)}{r_{31}(X_i - X_0) + r_{32}(Y_i - Y_0) + r_{33}(Z_i - Z_0)}(-f) \qquad 3.6$$

Notice that in equation $z_i = -f$ is used for the camera coordinates in the above equation.

The equations 3.6 are nonlinear equations, so the linearization is needed for solving them. If the number of control points are more than three, the least square method is needed. Refer to Ghost [1988] for the expression of $\boldsymbol{R}$ and details on solving the equation 3.6. Aerotriangulation is also available by software packages like Albany [ERIO Technologies].

With the development of the Global Positioning System (GPS), the exterior parameters of the camera can be precisely determined by the airborne differential GPS techniques or the combination of GPS and photogrammetry techniques.

3.2.2 Results of Aerotriangulation

Three different types of aerotriangulation methods are used here. They are: Albany from *ERIO*, Calib from Dr. *Jeyapalan*, my major professor, Softplotter from *Vision International*. The results from these three different software packages are available for the aerial photographs discussed in the chapter 2; and the results are shown in Table 3.1.

Table 3.1 Comparison of Triangulation Results Using the Three Different Software Packages

Parameter	Photograph No. 1			Photograph No. 2		
	Softplotter	Albany	Calib	Softplotter	Albany	Calib
X_0 (meters)	1487238.563	1487238.563	1487239.2576	1487792.068	1487791.822	1487791.865
Y_0 (meters)	1058385.185	1058385.689	1058386.2669	1058411.816	1058411.773	1058412.343
Z_0 (meters)	1229.815	1230.691	1203.4923*	1230.331	1230.980	1203.7212*
ω (degrees)	2.4992711	2.7455	2.1695047	2.4138072	2.6714	2.41656409
ϕ (degrees)	-2.6663514	-2.9648	-2.6786350	-2.7655344	-3.0804	-3.49479308
κ (degrees)	4.5477386	5.0931	4.2312360	2.7041475	7.1378	2.55292799

The heights with "*" are ellipsoid heights. The geoid undulation in Ames, state of Iowa area is about -29 meters, therefore the heights above mean sea level for those two exposure stations are close to 1230 meters. From the above table we see that the results are close, especially these results from Softplotter and Calib programs. Since Softplotter is used as a

comparison software in this research, the triangulation from Softplotter is used in the rectification of these two photographs.

3.2.3 Relative Orientation

The coplanerity condition must be met in a stereo pair. So the coplanerity condition can be used to check the stereo pair or to do the relative orientation as constraint. The relative orientation can also be used to extend the ground control for those photographs which do not have ground controls in the same strip by means of pass points. There are seven parameters involved in the relative orientation.

If the aerotriangulation is done with the ground control points and the pass points, the coplanerity condition is generally satisfied. The pass points are needed for checking the relative orientation if the exterior orientation parameters are imported from the results of other software packages like Albany.

3.2.4 Aerial Intersection for Locating Unknown Ground Points

After the exterior orientation is completed and the relative orientation is checked, equation 3.6 can be used to calculate the ground positions of any unknown points on the photographs in a stereo pair. Equation 3.6 is rewritten in the following format:

$$
\begin{aligned}
X &= \frac{r_{11}^{l}x_l + r_{21}^{l}y_l - r_{31}^{l}f_l}{r_{13}^{l}x_l + r_{23}^{l}y_l - r_{33}^{l}f_l}(Z - Z_0^{l}) + X_0^{l} \\
Y &= \frac{r_{12}^{l}x_l + r_{22}^{l}y_l - r_{32}^{l}f_l}{r_{13}^{l}x_l + r_{23}^{l}y_l - r_{33}^{l}f_l}(Z - Z_0^{l}) + Y_0^{l} \\
X &= \frac{r_{11}^{r}x_r + r_{21}^{r}y_r - r_{31}^{r}f_r}{r_{13}^{r}x_r + r_{23}^{r}y_r - r_{33}^{r}f_r}(Z - Z_0^{r}) + X_0^{r} \\
Y &= \frac{r_{12}^{r}x_r + r_{22}^{r}y_r - r_{32}^{r}f_r}{r_{13}^{r}x_r + r_{23}^{r}y_r - r_{33}^{r}f_r}(Z - Z_0^{r}) + Y_0^{r}
\end{aligned}
\qquad 3.7
$$

Where x_l, y_l, x_r, y_r are the camera coordinates for the same point on the left-hand and right-hand photographs in the stereo pair; (X_0^l, Y_0^l, Z_0^l) are the ground coordinates of the left-hand photograph; (X_0^r, Y_0^r, Z_0^r) are the ground coordinates of the right-hand photograph; $r_{11}^l, r_{12}^l, r_{13}^l, \cdots, r_{31}^l, r_{32}^l, r_{33}^l$ are the elements from the rotation matrix $\boldsymbol{R}$ for the left-hand photograph; $r_{11}^r, r_{12}^r, r_{13}^r, \cdots, r_{31}^r, r_{32}^r, r_{33}^r$ are the elements from the rotation matrix $\boldsymbol{R}$ for the right-

hand photograph; and f_l, f_r are the principal distances for the two photographs, respectively. The ground coordinates X, Y, Z of the unknown point are the three unknowns in equation 3.7. Since for each point on the ground there are four equations can be established, the ground position of each unknown point can be solved.

3.3 Single Photograph Rectification of Aerial Photographs

There are many ways to rectify the aerial photographs in photogrammetry. To do so in soft photogrammetry, digital single photograph rectification is used.

3.3.1 Process of Digital Rectification

Figure 3.4 shows the process of rectification. X, Y, Z and x, y, z are the ground coordinate and camera coordinate respectively; f is the focal length; X_0, Y_0, Z_0 are the ground coordinates of the camera nodal point **O**; the plane $a'b'$ is the photograph plane before rectification; the plane ab is the photograph plane after rectification; points A and B are two points on the ground and a, a', b, b' are their image points on the photographs before and after rectification, respectively; p' and p are the principal points on the photographs before and after rectification, respectively. In Figure 3.4 all the symbols with prime are related to the photograph before rectification.

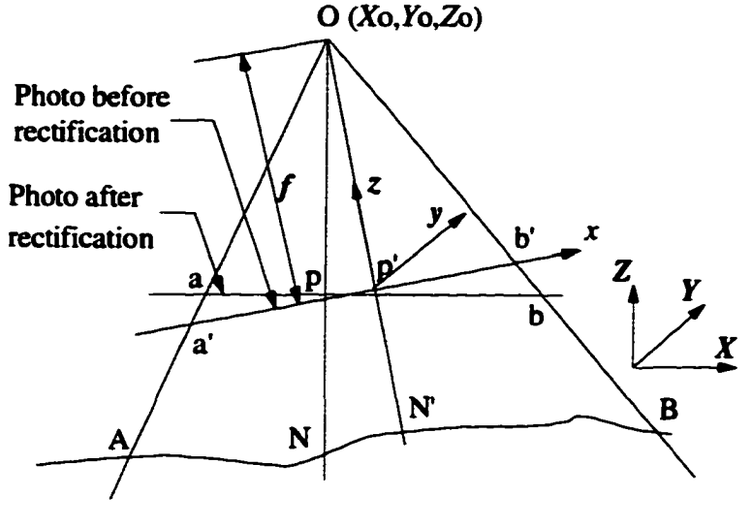


Figure 3.4 Rectification of Photographs

From the definition of angles and the order of the rotation from the ground coordinate system to the camera coordinate system, the aerial photographs can be brought back into the ground coordinate system by transforming the coordinate of each image pixel from the camera coordinate to the ground coordinate or the rectified camera coordinate. So the points a', b' will be transformed to points a, b, respectively on the photograph after the single rectification.

3.3.2 Single Rectification Equation

Combine equations 3.2 and 3.6 to obtain the rectification equation as:

$$x_r = \frac{r_{11}x_c + r_{21}y_c - r_{31}f}{r_{13}x_c + r_{23}y_c - r_{33}f}(-f)$$
$$y_r = \frac{r_{12}x_c + r_{22}y_c - r_{32}f}{r_{13}x_c + r_{23}y_c - r_{33}f}(-f) \qquad 3.8$$

Where z_r is replaced by $-f$ for the same scale rectification. The pixel gray values must be used when displaying the digital photograph, the relation between the pixel and camera, rectified coordinates is determined by the position of the initial point and the pixel size by using the following equation:

$$x_c = x_{0c} + C_{xc}I_c$$
$$y_c = y_{0c} + C_{yc}J_c$$
$$x_r = x_{0r} + C_{xr}I_r \qquad 3.9$$
$$y_r = y_{0r} + C_{yr}J_r$$

Where I_c, J_c, I_r, J_r are the pixel coordinates for the camera and rectified coordinates respectively; C_{xc}, C_{yc}, C_{xr}, C_{yr} are the pixel sizes for the camera and rectified coordinates, respectively; x_c, y_c, x_r, y_r are the coordinates of the initial points of the camera and rectified coordinates respectively; and (x_{0c}, y_{0c}); (x_{0r}, y_{0r}) are the initial positions for the camera and rectified coordinate systems, respectively.

3.3.3 Determination of Initial Point, Pixel Size and Image Size of the Rectified Photographs

3.3.3.1 Pixel Size

The same pixel size as the photographs before rectification will be used for the photographs after rectification. Since the shape of the pixels is not square, the direction relationship between the pixel before and after rectification should be considered. The pixel

size in x and y directions should be switched if the coordinate system is rotated more than 45°. In equation assume the coordinates of a point in the camera coordinate system is (1, 0, -f), then the rectified coordinates of the point are:

$$x_r = \frac{r_{11} - r_{31} f}{r_{13} - r_{33} f}(-f)$$

$$y_r = \frac{r_{12} - r_{32} f}{r_{13} - r_{33} f}(-f)$$

The values of x_r and y_r can be used to determine if the switch is needed as in equation 3.10

$$\begin{aligned} &\textit{if} \quad |x_r| < |y_r| \quad \textit{then} \\ &\quad C_{xr} = C_{xc} \\ &\quad C_{yr} = C_{yc} \\ &\textit{else} \\ &\quad C_{xr} = |C_{yc}| \\ &\quad C_{yr} = \frac{C_{xc} C_{yc}}{|C_{yc}|} \end{aligned} \qquad 3.10$$

The absolute symbol used in the above equation ensures that the signs of the values are correct if the switch is performed.

3.3.3.2 Determination of Initial Points and Image Size

From equation 3.8 the rectified coordinates of the four corners on the rectified photo can be determined by substituting the camera coordinates of the four corners of the photograph before rectification. Assume the four corners of the rectified photograph have the maximum and minimum values of x_{min}, x_{max}, y_{min} and y_{max} from equation 3.8, then the initial point location and the image size are given by equation 3.11.

In equation 3.11 $size\_I_r$, $size\_J_r$ are the pixel size of the rectified image. Other symbols are the same as in the previous section.

$$
\begin{aligned}
x_{0r} &= x_{\min} \\
y_{0r} &= y_{\min} \; if \quad C_{yr} > 0 \\
y_{0r} &= y_{\max} \; if \quad C_{yr} < 0 \\
size\_I_r &= \frac{x_{\max} - x_{\min}}{C_{xr}} \\
size\_J_r &= \frac{y_{\max} - y_{\min}}{|C_{yr}|}
\end{aligned}
\qquad 3.11
$$

3.3.4 Resampling Equation in Rectification of Photographs

Same as the transformation in the interior process, digital resampling is necessary in the process of rectification. The actual formula used in rectification is the resampling formula. The single rectification resampling formula needed is the reverse form of equation 3.8, which is similar to equation 3.6. Combine equations 3.6 and 3.9 to obtain the resampling equation as

$$
\begin{aligned}
A_1 &= r_{11}x_{0r} + r_{12}y_{0r} - r_{13}f \\
B_1 &= r_{21}x_{0r} + r_{22}y_{0r} - r_{23}f \\
C_1 &= r_{31}x_{0r} + r_{32}y_{0r} - r_{33}f \qquad 3.12 \\
I' &= -\frac{(z_r r_{11} + x_{0c} r_{31})C_{xr} I_r + (z_r r_{12} + x_{0c} r_{32})C_{yr} J_r + (z_r A_1 + x_{0c} C_1)}{C_x\left[r_{31}C_{xr} I_r + r_{32}C_{yr} J_r + (r_{31}x_{0r} + r_{32}y_{0r} - r_{33}f)\right]} \\
J' &= -\frac{(z_r r_{21} + y_{0c} r_{31})C_{xr} I_r + (z_r r_{22} + y_{0c} r_{32})C_{yr} J_r + (z_r B_1 + y_{0c} C_1)}{C_y\left[r_{31}C_{xr} I_r + r_{32}C_{yr} J_r + (r_{31}x_{0r} + r_{32}y_{0r} - r_{33}f)\right]}
\end{aligned}
$$

Where A_1, B_1, C_1 are three intermediate symbols; I', J' are the camera coordinates calculated from the given rectified coordinates I_r, J_r which range from 0, 0 to $size\_I_r$-1, $size\_J_r$-1 (see equation 3.11) respectively; z_r is the z coordinate in the rectified coordinate system and here is set to $-f$. Since I', J' may not be integers, the interpolation methods discussed in Chapter 2 are applied. To simplify the computation in the rectification, rewrite the equation 3.12 as equation 3.13.

Where k is a constant; M is called rectification matrix; A_1, B_1 are the same constants as that in equation 3.12. So the rectification can be done at any scale.

$$\begin{bmatrix} I' \\ J' \\ -f \end{bmatrix} = M \begin{bmatrix} I_r \\ J_r \\ -f \end{bmatrix} k$$

$$M = \begin{bmatrix} \frac{(z_r r_{11} + x_{0c} r_{31}) C_{xr}}{-C_{xc}} & \frac{(z_r r_{12} + x_{0c} r_{32}) C_{yr}}{-C_{xc}} & \frac{z_r A_1 + x_{0c} C_1}{-C_{xc}} \\ \frac{(z_r r_{21} + y_{0c} r_{31}) C_{xr}}{-C_{yc}} & \frac{(z_r r_{22} + y_{0c} r_{31}) C_{xr}}{-C_{yc}} & \frac{z_r B_1 + y_{0c} C_1}{-C_{yc}} \\ r_{31} C_{xr} & r_{32} C_{yr} & r_{31} x_{0r} + r_{32} y_{0r} - r_{33} f \end{bmatrix} \quad 3.13$$

3.4 Results of the Rectification of Aerial Photographs

The rectification results of these two aerial photographs are represented in this section.

3.4.1 LAN Headers of the Rectified Photographs

Table 3.2 shows some of contents of the headers of these two rectified photographs. Notice that different item may have different type of unit.

Table 3.2 Content of Some Items of the LAN Header of the Rectified Photographs

Name	Photograph No. 1	Photograph No. 2
X_0 (meters)	1487238.563	1487792.068
Y_0 (meters)	1058385.185	1058411.816
Z_0 (meters)	1229.815	1230.331
x_{0c} (mm)	-15.343071	-103.299477
y_{0c} (mm)	129.266998	117.976028
f (mm)	152.473	152.473
size_x_c	2940	2865
size_y_c	4029	3833
C_{xc} (mm)	0.049580	0.049370
C_{yc} (mm)	-0.049521	-0.049239

Comparing this table with Table 2.5 it is known that the pixel sizes remain the same, but the initial positions are different. The image sizes are also changed because the rotation in the rectification process. The other differences are that the location for each photographs in the ground coordinate system and the focal length are added to the header.

3.4.2 Images of the Rectified Photographs

These two rectified aerial photographs are shown in Figures 3.5 and 3.6. Comparing them with the two photographs before rectification shown in Figures 2.9 and 2.10, the differences between them are obvious. Since the roads are in due east and north directions, the rectified images look more correct.

3.5 Conclusions

Single photograph rectification, triangulation, and resampling are discussed in this chapter. Different types of interpolation methods are used. The transformation and resampling equations for the single photograph rectification are derived, the parameters in the transformation are also determined. All these equations and determination of parameters are implemented in computer programs. Two photographs and some information in their header after the single photograph rectification are shown and examined.

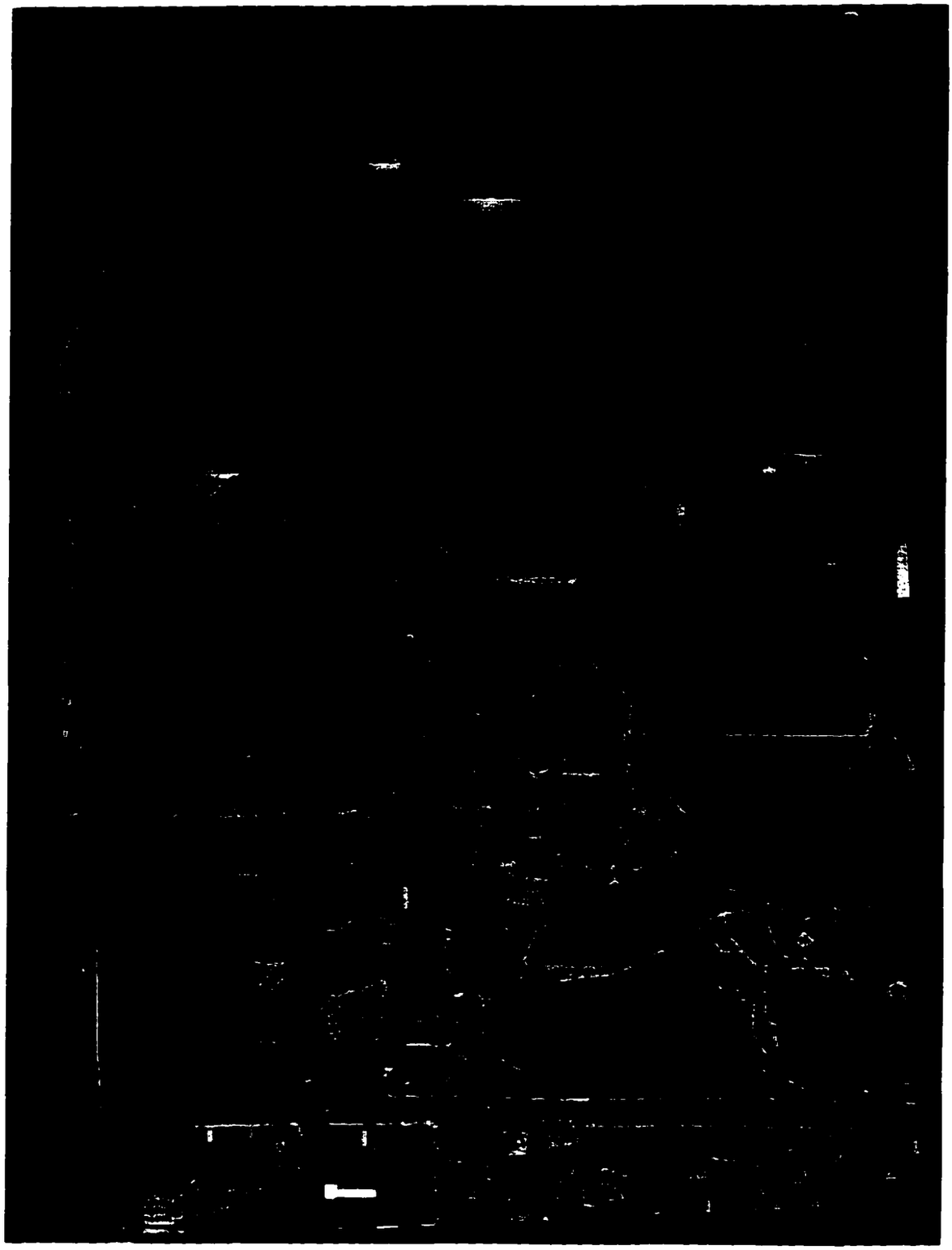


Figure 3.5 Rectified Photograph from Photograph No. 1

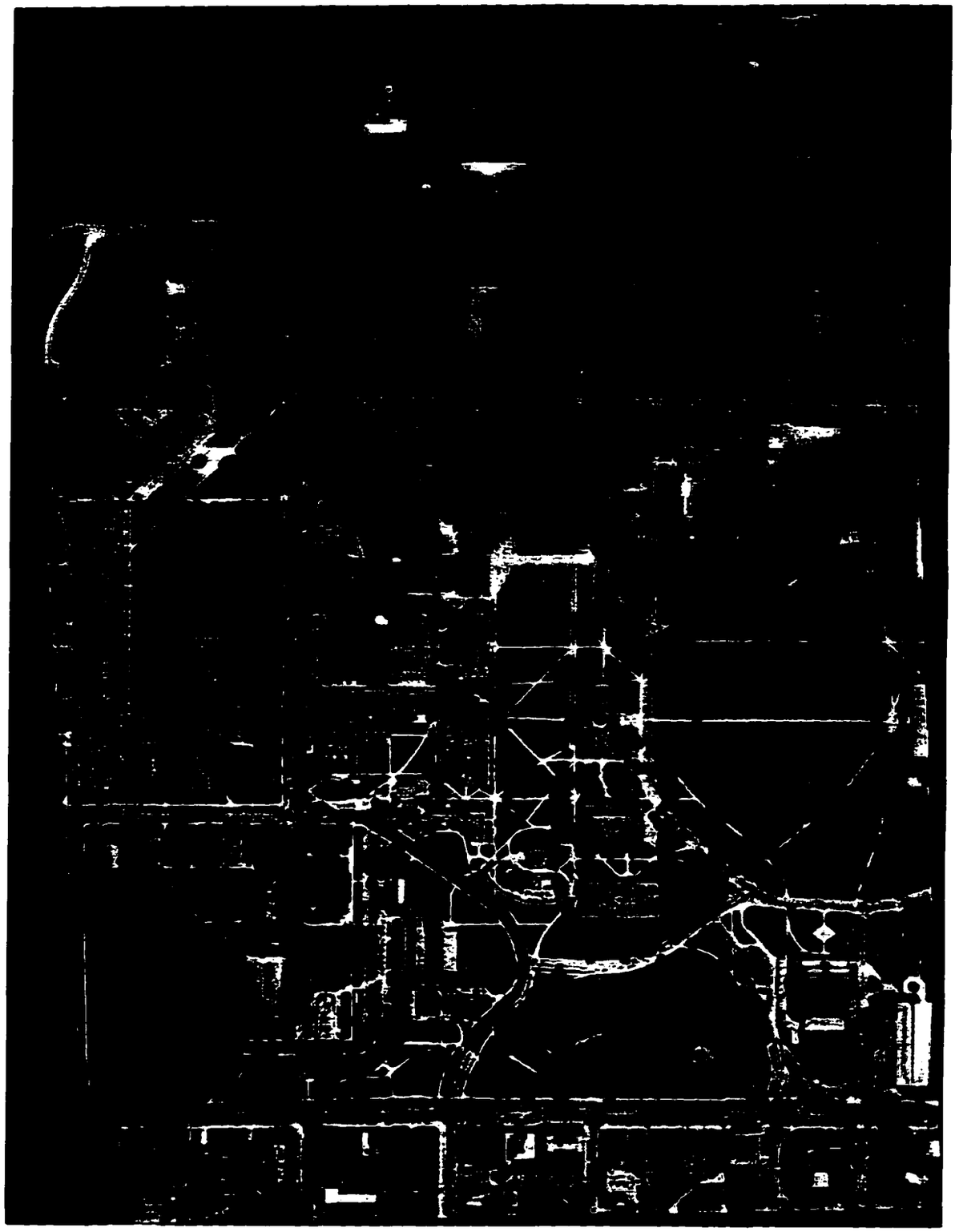


Figure 3.6 Rectified Photograph from Photograph No. 2

CHAPTER 4 STEREO RECTIFICATION

In Chapter 3 the rectification of single photograph is described. In a stereo pair two or more photographs that are corrected. Before the image matching can be performed, all the photographs in the same stereo pair must have the same pixel size, scale and be in the same coordinate system - the epipolar coordinate system. This processing is called stereo rectification.

4.1 Epipolar Coordinate System

As shown in Figure 4.1, the direction of x_e axis of the epipolar coordinate system is defined by the line O'O" between the nodal points of the adjacent cameras of the photographs in the stereo pair. In Figure 4.1, $x'y'z', x''y''z''$ are the rectified coordinates; $x'_e y'_e z'_e, x''_e y''_e z''_e$ are the epipolar coordinates. Each photograph has its own epipolar coordinate system in the stereo pair. The only difference between the epipolar coordinate systems is the different origins.

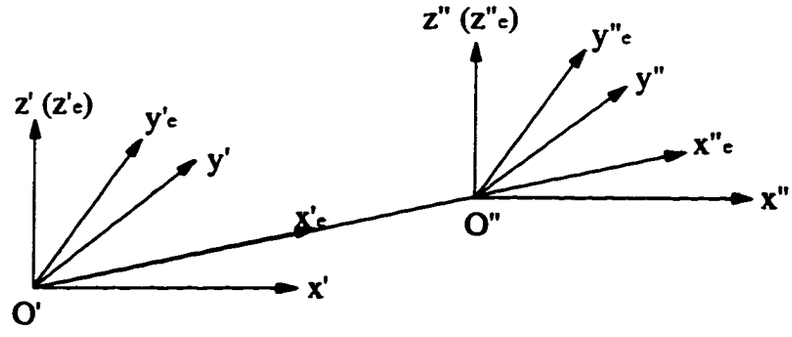


Figure 4.1 Epipoplar Coordinate System

To transform from rectified coordinates to epipolar coordinates, only rotation around the z axis is involved. So the origin and z axis of the rectified coordinate are not changed in this transformation provided the scales are the same.

4.2 Equalizing Pixel Size in A Stereo Pair

The pixel sizes of the photographs in the same stereo pair are assumed to be the same in the directions of both rows and columns to make the image matching simple. Thus the number of the pixels can be used as the unit of distance on each photograph in the stereo pair without concern for real distance provided the scales of the photographs are the same.

4.2.1 Image Resizing

Pixel size is changed in the process of image resizing. Although the pixel coordinates of any location on the photograph could be different, the rectified coordinates of any location on the photograph will remains the same in this process. Assume the values of pixel size in columns and rows are C_x, C_y and C'_x, C'_y before and after resizing, respectively and I, J and I', J' are the number of columns and rows of the photograph before and after resizing, respectively. Since the distance remains the same in the resizing and thus the coordinates of the initial pixel are the same for both photographs, the transformation equation is:

$$I' = \frac{C_x}{C'_x} I$$
$$J' = \frac{C_y}{C'_y} J \qquad 4.1$$

The resized photograph has more columns and rows if its pixel size is smaller than that of the photograph before resizing. Since the coordinates of the initial pixel in the upper-left corner are not changed, the coordinates of the initial position are not needed in equation 4.1. The values of C_x / C'_x and C_y / C'_y here are the scale factors. If the scale factor for a direction is larger than 1, then the image after resizing will look larger with smaller pixel size in this direction.

4.2.2 Size of the Resized Photograph

The size of the resized photograph can be determined by equation 4.1 as:

$$size\_x' = \frac{size\_xC_x}{C_x'}$$
$$size\_y' = \frac{size\_yC_y}{C_y'} \qquad 4.2$$

Where ***size_x***, ***size_y*** and ***size_x'***, ***size_y'*** are the size of the photograph in number of columns and rows before and after resizing, respectively.

4.2.3 Resampling in Photograph Resizing

The resampling formula is:

$$I = \frac{C_x'}{C_x} I'$$
$$J = \frac{C_y'}{C_y} J' \qquad 4.3$$
$$I' = 0, \cdots, size\_x' - 1$$
$$J' = 0, \cdots, size\_y' - 1$$

Since ***I***, ***J*** could be floating numbers, the interpolation mentioned before can be used to determine the pixel gray value on the resized photograph for all the pixels on that photograph.

4.2.4 Image Resizing in A Stereo Pair

One photograph in the stereo pair is used as the reference photograph in this process. Its pixel size is used for both photographs in this stereo pair. After resizing, all the photographs in this stereo pair should have the same pixel size in both ***x*** and ***y*** directions.

Table 4.1 Some Information in the Header after Resizing

Name	Photograph No. 2	
	Before Resizing	After Resizing
x_{0r} (mm)	-103.299477	-103.299477
y_{0r} (mm)	117.976028	117.976028
Cxr (mm)	0.049370	0.049580
C_{yr} (mm)	-0.049239	-0.049521
$size\_x_r$	2865	2853
$size\_y_r$	3833	3811

Here, photograph No. 1 is considered as the reference photograph in the stereo pair. After resizing this stereo pair, some of the information in the header of photograph No. 2 are shown in Table 4.1. In order to make the comparison between the photographs before and after resizing, the header information of both photographs are shown in Table 4.1. According to the listed information, the differences between them are the image size and the pixel size. The image after resizing looks the same as the image before resizing since there is no change in the rectified coordinates of points on the image. The information about the photograph No. 1 is not shown in Table 4.1 since it is the reference image. Comparing Table 4.1 with Table 3.3 it can be seen that the two photographs have the same pixel size in both x and y directions.

4.3 Determination of the Scale of the Stereo Pair

The photographs in the same stereo pair must have the same scale before image matching. By equalizing the scale in both photographs the pixels would have the same ground size at the average ground height of the stereo model.

4.3.1 Scale of A Photograph

The scale of each photograph is the function of the flight height, principal distance, and the ground height expressed as:

$$s = \frac{f}{Z_0 - h_a} \qquad 4.4$$

for the vertical or rectified photograph. Where Z_0 is the flight height above mean sea level (MSL); h_a is the average height of the ground above MSL; f is the principal distance or the z_r distance after the single photograph rectification. Since the ground height is different from point to point on the photograph and not available at this time, the average ground height is used in equation 4.4. Because there is only one common area involved in one stereo pair, the average height is the same for all the photographs in the stereo pair.

In equation 4.4, the principal distance for the photographs taken by the same camera is the same. If a photograph in the stereo pair was taken by a different camera at different object distance, its principal distance may be different from the other.

To change the scale of a photograph in the stereo pair, one can change the flight height or the principal distance. Since different flight height is the most likely reason to cause the different scales in the stereo pair, changing flight height method is used. In this a way, the principal distance can remain the same.

4.3.2 Determination of the New Flight Height

To get an equal scale in the stereo pair, changing flight height is used to change the photograph scale. One photograph in the stereo pair is used as the reference photograph. Assume its flight height and principal distance are H_r and f_r respectively, the changed photographs have H_o, H_n as old and new flight heights, respectively; and f as the principal distance. To have equal scale,

$$\frac{f}{H_n - h_a} = \frac{f_r}{H_r - h_a}$$

is needed. Therefore,

$$H_n = \frac{(H_r - h_a)f}{f_r} + h_a \qquad 4.5$$

If the photograph has the same principal distance as the reference photograph, equation 4.5 becomes

$$H_n = H_r$$

Since the photographs in the stereo pair are after single photograph rectification, the f here refers to the principal distance which is the same as the z_r in equation 3.2. The principal distance in the photographs after single photograph rectification could be different from the camera focal length, it depends on the single photograph rectification scale of the each photograph.

4.3.3 Transformation between the Old and New Coordinates

Assume the s is the ratio of the new scale s_n and the old scale s_o, called scale unification factor, the relation between them is:

$$s = \frac{s_n}{s_o} = \frac{f_r(H_o - h_a)}{f(H_r - h_a)} = \frac{(H_o - h_a)}{(H_n - h_a)} = \frac{x_n}{x_o} = \frac{y_n}{y_o} \qquad 4.6$$

When $f_r = f$. Where x_o, y_o and x_n, y_n are the rectified coordinates for each pixel on the photographs before and after the transformation, respectively. s is the magnification factor for the photographs to be changed, and equation 4.6 is the transformation equation for the equalizing scale process.

4.3.4 Determination of New Initial Position and Image Size

From equation 4.6 the relation between the old and new initial position is

$$\begin{aligned} x'_{0r} &= sx_{0r} \\ y'_{0r} &= sy_{0r} \end{aligned} \qquad 4.7$$

and the new photograph size is

$$\begin{aligned} size\_x_{nr} &= size\_x_{or}s \\ size\_y_{nr} &= size\_y_{or}s \end{aligned} \qquad 4.8$$

Where x'_{0r}, y'_{0r}, $size\_x_{nr}$, $size\_y_{nr}$ are for the new photograph after this process; $size\_x_{or}$, $size\_y_{or}$ are for the old photograph. Unlike in section 4.2, here the coordinates of the initial position do not remain the same after the transformation.

4.3.5 Transformation Equation

Since the equalizing scale process is after the image resizing discussed in the previous section, the photographs in the stereo pair should have the same pixel size, and consider equation 4.6, then transformation equation in pixel coordinate is

$$\begin{aligned} I_{nr} &= sI_{or} \\ J_{nr} &= sJ_{or} \end{aligned} \qquad 4.9$$

Where I_{or}, J_{or} and I_{nr}, J_{nr} are the pixel coordinates before and after transformation, respectively. Notice that the pixel size remain the same after transformation.

4.3.6 Resampling Formula

The resampling formula is the reverse process of equation 4.9. It is given by

$$\begin{aligned} I'_{or} &= \frac{I_{nr}}{s} \quad I_{nr} = 0, \cdots, size\_x_{nr} - 1 \\ J'_{or} &= \frac{J_{nr}}{s} \quad J_{nr} = 0, \cdots, size\_y_{nr} - 1 \end{aligned} \qquad 4.10$$

One of these interpolation methods can be applied to complete the transformation.

4.3.7 Results of Equalizing Scale

As before, photograph No. 1 is considered as the reference photograph. Table 4.2 shows some new information about the LAN header of photograph No. 2 after the transformation. Since the scale difference between these two photographs is small, the change of image size for the photograph No. 2 is small. In Table 4.2 the ground height is assumed as 0. The new photograph scale is larger then the old one because s is larger than 1. So the initial pixel is farther to the coordinate origin than it was before. Notice that the flight height in the header remains the same as before. s can give the new height by using equation 4.6. The pixel size remains the same so it is not shown in Table 4.2. The information of photograph No. 1 is not shown in Table 4.2 since it is unchanged.

Table 4.2 Some Information about the Header of Photograph No. 2

Name	Photograph No. 2	
	Before Scaling	After Scaling
s	1	1.000420
x_{0r} (mm)	-103.299477	-103.342827
y_{0r} (mm)	117.976028	117.025536
Z_0 (meters)	1230.331	1230.331
$size\_x_r$	2853	2854
$size\_y_r$	3811	3812

The s in Table 4.2 is defined in equation 4.6. Comparing Table 4.3 with Table 3.3 it can be seen that the flight height remains the same. According to equation 4.6 the flight height after equalizing the scale is

$$H = \frac{(Z_0 - h_a)}{s} + h_a$$

therefore the value of H from the above equation and the Z_0 from Table 4.2 is about 1229.815 Meters which is equal to the flight height of the photograph No. 1 in Table 3.3.

Figure 4.2 shows the photograph No. 2 in the stereo pair after equalizing the scale.

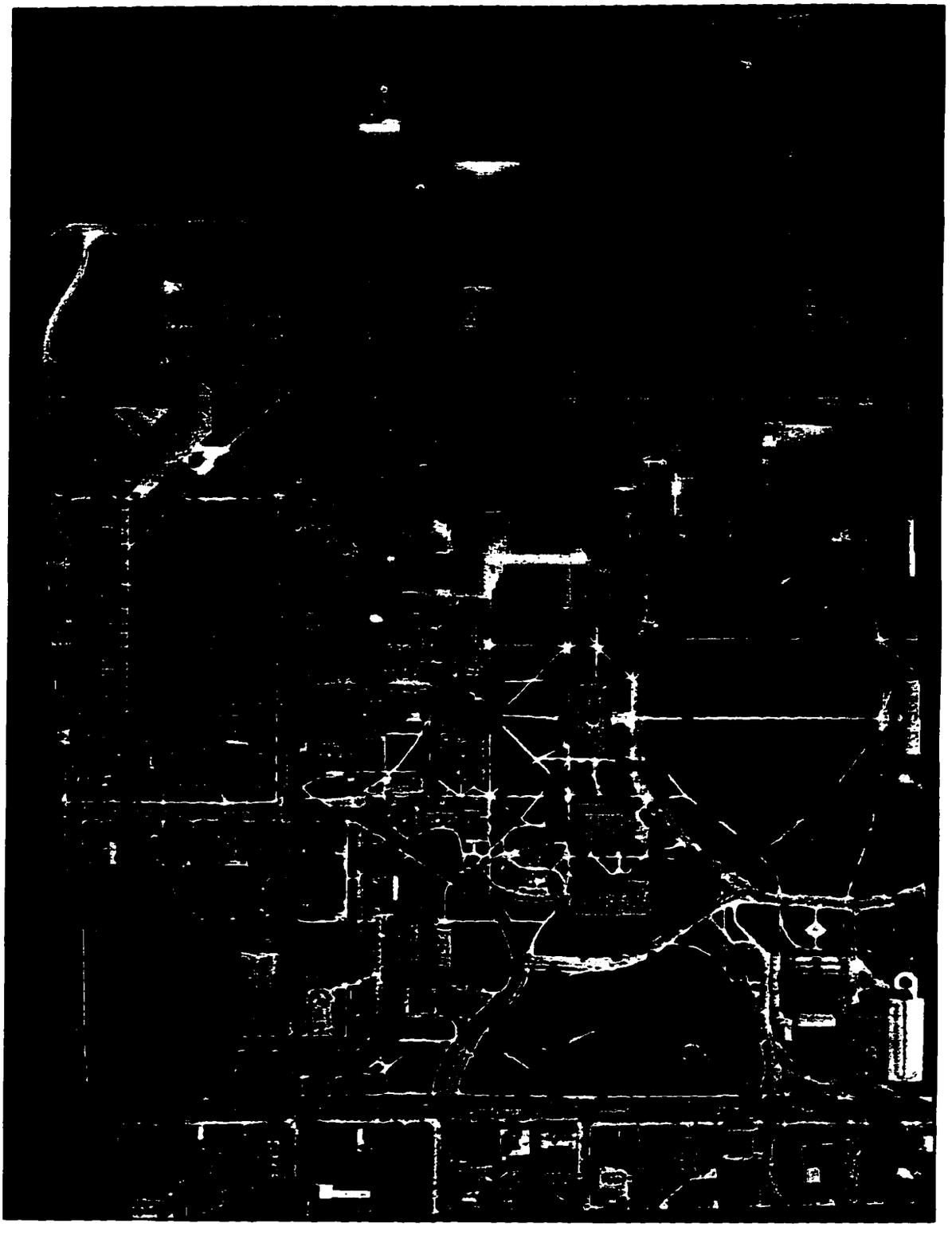


Figure 4.2 Photograph No. 2 after Equalizing Scale

4.3.8 Error in the Equalizing Scale Process and the Correction

In the equalizing scale process, the average ground height is used to determine the scale of each photograph. But the actual scale on the photograph varies from point to point. Therefore, position error will occur on the photograph for those pixels whose ground heights are not equal to the average ground height after processing.

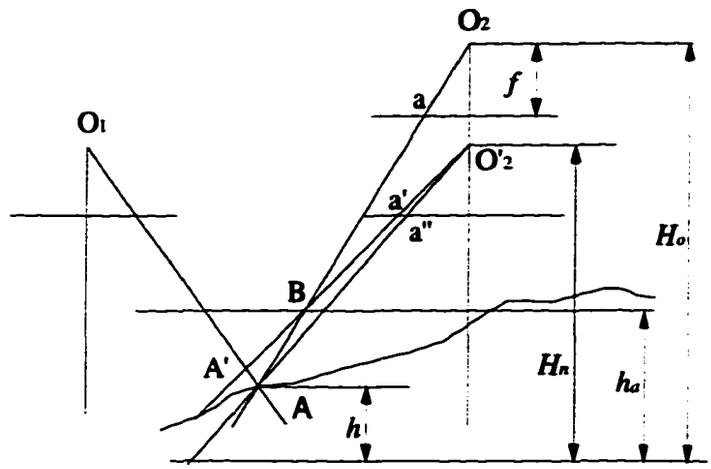


Figure 4.3 Error Appears in Equalizing Scale Process

As shown in Figure 4.3, O_2, O'_2 are the locations for the right-handed photograph before and after the processing, respectively; point A is the point on the ground; point a is the image of the point A on the right-handed photograph before the processing. Point B in the intersection of line AO_2 and the average ground height h_a. Since the scaling is performed according to the average ground height h_a, the image point of the point A on the processed photograph is located on point a', but its correct position should be located on point a", according to the geometric relationship shown in Figure 4.3. Therefore the error of distance a'a" is introduced after the processing. If the position a' is used to compute the height of the point A after image matching, if will give the wrong height information. So correction is needed when computing the ground height.

In Figure 4.3 the relationship between point a and a' is determined by the scaling factor s as $d_{a'} = s d_a$.

Where d_a, $d_{a'}$ are the radial distances of the image point of the ground point A in the rectified coordinates before and after equalizing the stereo pair scale.

If the matched point a' is converted back to the original photograph with the original height, according to the geometry shown in Figure 4.3, the correct height of the point A is available. Therefore, before calculating the location for each pixel after image matching, the pixel position should be corrected by dividing the scaling factor s. In this way the error brought by the scale equalizing process can be removed when computing the ground height.

4.4 Epipolar Rectification

After all the previous processing are done, all the photographs in the stereo pair should have same photo scale and flight height, then the epipolar rectification can be performed. In the process of transforming to the epipolar coordinate system, only one rotation around z axis of the rectified coordinate needs to be done.

4.4.1 Determination of the Rotation Angle for Epipolar Rectification

The location of each camera is known after the photograph has been rectified. Assume the locations of the photographs are (X_l, Y_l), (X_r, Y_r) in the ground coordinate system for the left and right photographs shown in Figure 4.5 (Z axes are not shown).

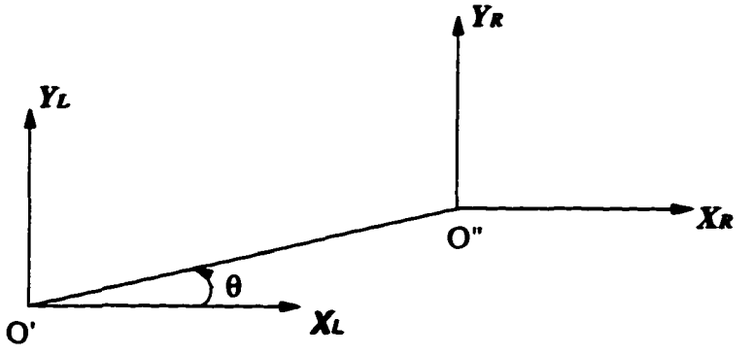


Figure 4.4 Determination of Rotation Angle θ

The rotation angle is given by:

$$\theta = \tan^{-1}\left(\frac{Y_r - Y_l}{X_r - X_l}\right) \qquad 4.11$$

The rectified coordinate is transformed to epipolar coordinate system by rotating angle θ around z_r axis. The Z axes are not shown in Figure 4.4. The sign of the angle θ is positive when the rotation is counter-clockwise.

4.4.2 Transformation Formula

The transformation equation is given by equation 4.12. Where x_e, y_e and x_r, y_r are the epipolar and rectified coordinates, respectively; S is the epipolar transformation matrix.

$$\begin{bmatrix} x_e \\ y_e \\ z_e \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = S \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} \qquad 4.12$$

Using the pixel coordinates to represent the equation 4.12 as:

$$\begin{bmatrix} I_e \\ J_e \end{bmatrix} = \begin{bmatrix} \frac{\cos\theta}{C_{xe}} & \frac{\sin\theta}{C_{xe}} \\ \frac{-\sin\theta}{C_{ye}} & \frac{\cos\theta}{C_{ye}} \end{bmatrix} \begin{bmatrix} C_{xr} I_r \\ C_{yr} J_r \end{bmatrix} + \begin{bmatrix} \frac{x_{0r}\cos\theta + y_{0r}\sin\theta - x_{0e}}{C_{xe}} \\ \frac{-x_{0r}\sin\theta + y_{0r}\cos\theta - y_{0e}}{C_{ye}} \end{bmatrix} \qquad 4.13$$

Where I_e, J_e and I_r, J_r are the pixel coordinates for the epipolar and rectified coordinates, respectively; x_{0e}, y_{0e} and x_{0r}, y_{0r} are the initial locations for epipolar and rectified coordinates, respectively; C_{xe}, C_{ye} and C_{xr}, C_{yr} are the pixel sizes for epipolar and rectified coordinates, respectively.

4.4.3 Determination of the New Image Size, Initial Location and the Pixel Sizes

The equation 4.12 can be used to determine the initial location of the new photograph. Assume the image sizes for both photographs are ***size_x***$_e$, ***size_y***$_e$ and ***size_x***$_r$, ***size_y***$_r$, the new image limit can be determined by using equation 4.12, say x_{min}, x_{max}, y_{min}, y_{max}, then image constants for the new image size are given by equation 4.14.

After the epipolar transformation, the stereo pair is in the epipolar coordinate system. If the stereo rectification is perfect, there will be no *y* parallax in the stereo model. A good relative orientation may help to reduce the *y* parallax in the stereo model significantly.

$$
\begin{aligned}
&\textit{if} \quad |\theta| < 45^\circ \quad \textit{then} \\
&C_{xe} = C_{xr} \\
&C_{ye} = C_{yr} \\
&\textit{else} \\
&C_{xe} = |C_{yr}| \\
&C_{ye} = C_{xe} \frac{C_{yrr}}{|C_{yr}|} \\
&size\_x_e = \frac{x_{\max} - x_{\min}}{C_{xe}} \\
&size\_y_e = \frac{y_{\max} - y_{\min}}{|C_{ye}|} \\
&x_{0e} = x_{\min} \\
&y_{0e} = y_{\min} \textit{ if } \quad C_{ye} > 0 \\
&y_{0e} = y_{\max} \textit{ if } \quad C_{ye} < 0
\end{aligned}
\qquad 4.14
$$

4.4.4 Resampling Formula

Like the processes mentioned before, resampling is needed for performing the digital image transformation. Reversing the equation 4.13 to get the resampling formula as:

$$
\begin{bmatrix} I_r \\ J_r \end{bmatrix} = \begin{bmatrix} \frac{\cos\theta}{C_{xr}} & \frac{-\sin\theta}{C_{xr}} \\ \frac{\sin\theta}{C_{yr}} & \frac{\cos\theta}{C_{yr}} \end{bmatrix} \begin{bmatrix} C_{xe} I_e \\ C_{ye} J_e \end{bmatrix} + \begin{bmatrix} \frac{x_{0e}\cos\theta - y_{0e}\sin\theta - x_{0r}}{C_{xr}} \\ \frac{x_{0e}\sin\theta + y_{0e}\cos\theta - y_{0r}}{C_{yr}} \end{bmatrix}
$$

$$
\begin{aligned}
&I_e = 0, \cdots, size\_I_e; \\
&J_e = 0, \cdots, size\_J_e;
\end{aligned}
\qquad 4.15
$$

Interpolation will be applied.

4.4.5 Results of Epipolar Rectification

Table 4.3 shows part of the LAN header of the photographs after epipolar rectification. The θ in Table 4.3 is the angle rotated in the process of epipolar rectification. The photographs after epipolar rectification are shown in Figures 4.5 and 4.6.

Table 4.3 Part of the Header Information after Epipolar Rectification

Name	Photograph No. 1	Photograph No. 2
x_{0r} (mm)	-18.699175	-106.621063
y_{0r} (mm)	129.854996	122.855591
Z_0 (meters)	1229.815	1230.331
s	1	1.000420
θ (degrees)	2.754570	2.754570
C_{xr} (mm)	0.049580	0.049580
C_{yr} (mm)	-0.049521	-0.049521
$size\_x_r$	3130	3034
$size\_y_r$	4166	3945

Since some areas on the photographs after transformation show no features because of the rotations, it is better to redefine the common area to be used in image matching. The new common area is shown in Figures 4.7 and 4.8 for both images; and Table 4.4 shows the new LAN header information for the new common area on each photograph.

Table 4.4 Some of the Header Information about the New Common Area

Name	Photograph No. 1	Photograph No. 2
x_{0r} (mm)	-9.576458	-97.894989
y_{0r} (mm)	123.516304	116.714989
Z_0 (meters)	1229.815	1230.331
s	1	1.000420
θ (degrees)	2.754570	2.754570
C_{xr} (mm)	0.049580	0.049580
C_{yr} (mm)	-0.049521	-0.049521
$size\_x_r$	2753	2689
$size\_y_r$	3933	3772

4.5 Comparison of Rectification with SoftPlotter

The results from rectification of a single photograph and stereo rectification are compared with that from the SoftPlotter. For the comparison in the rest of the study, 24 points are selected and their locations are manually determined. The heights determined are also calculated and compared in this section.

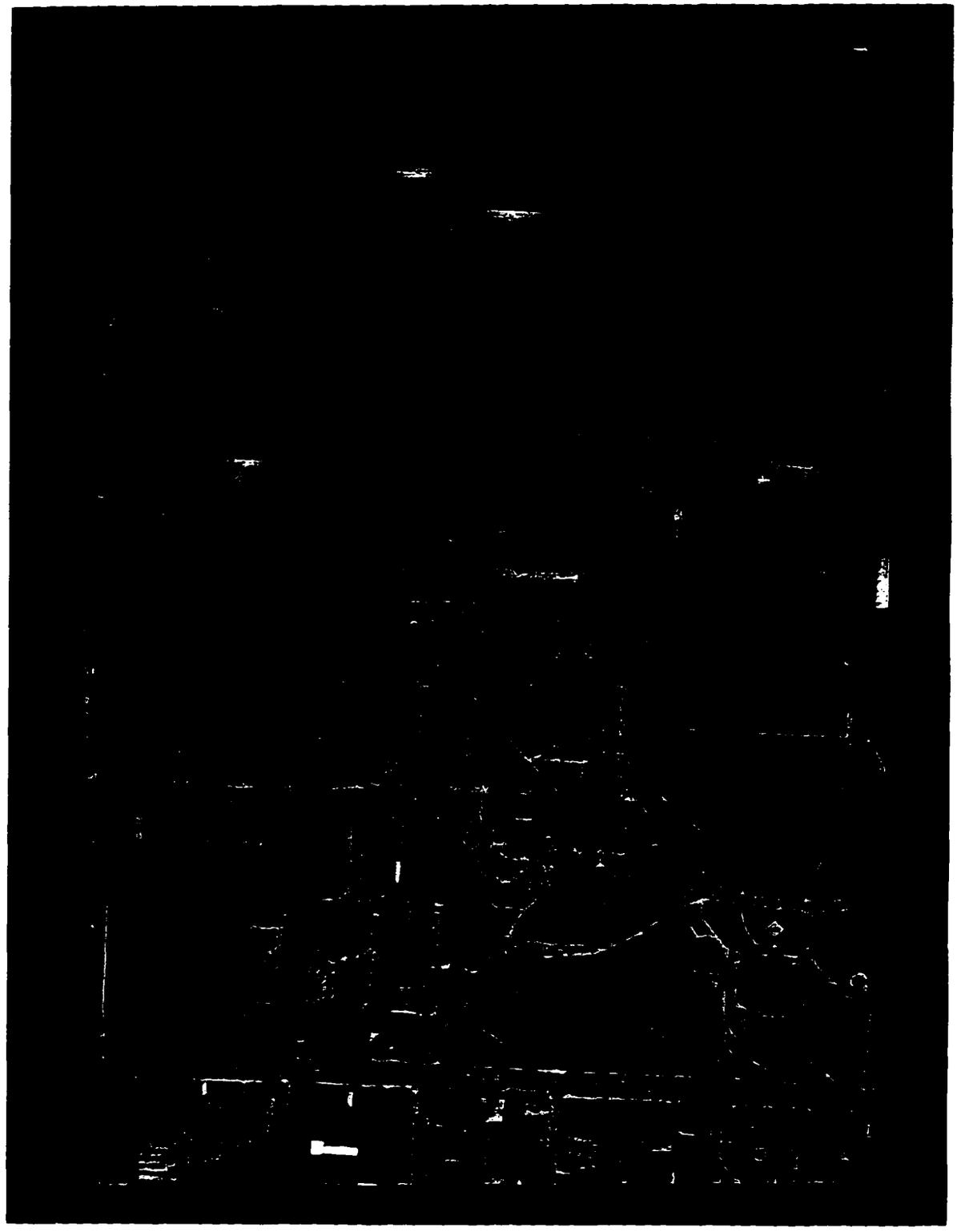


Figure 4.5 Photograph No. 1 after Epipolar Rectification

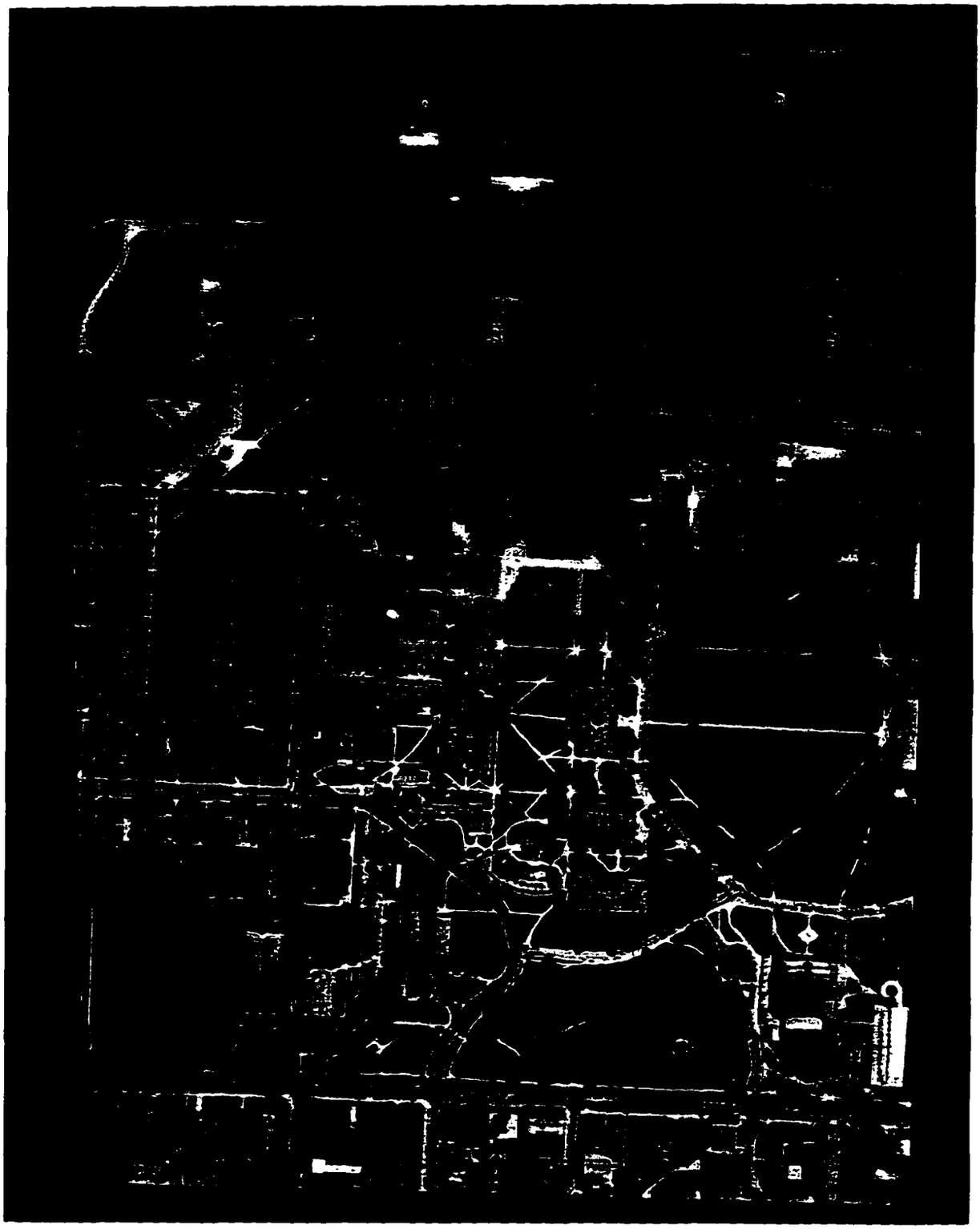


Figure 4.6 Photograph No. 2 after Epipolar Rectification

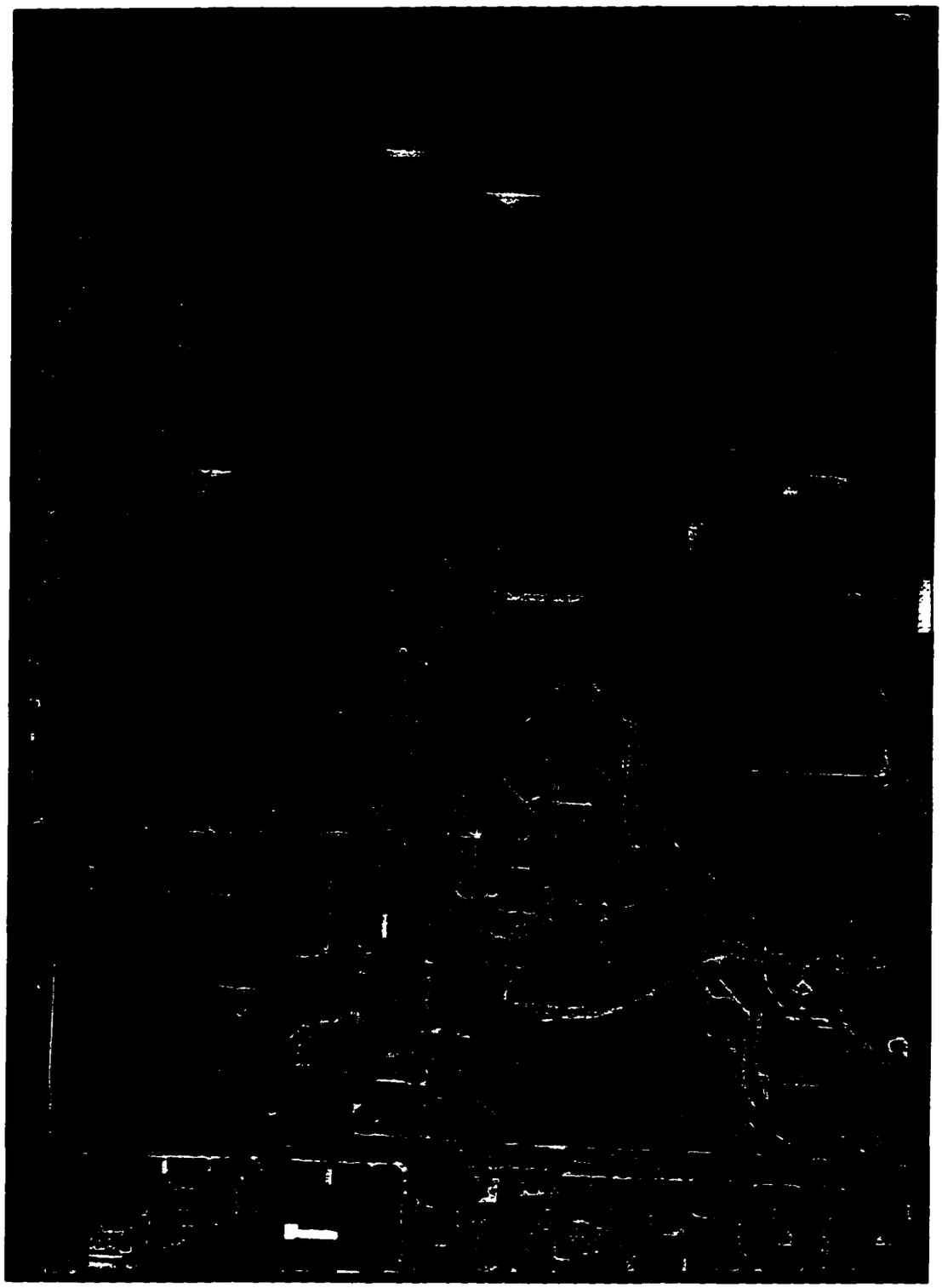


Figure 4.7 New Common Area on Photograph No. 1

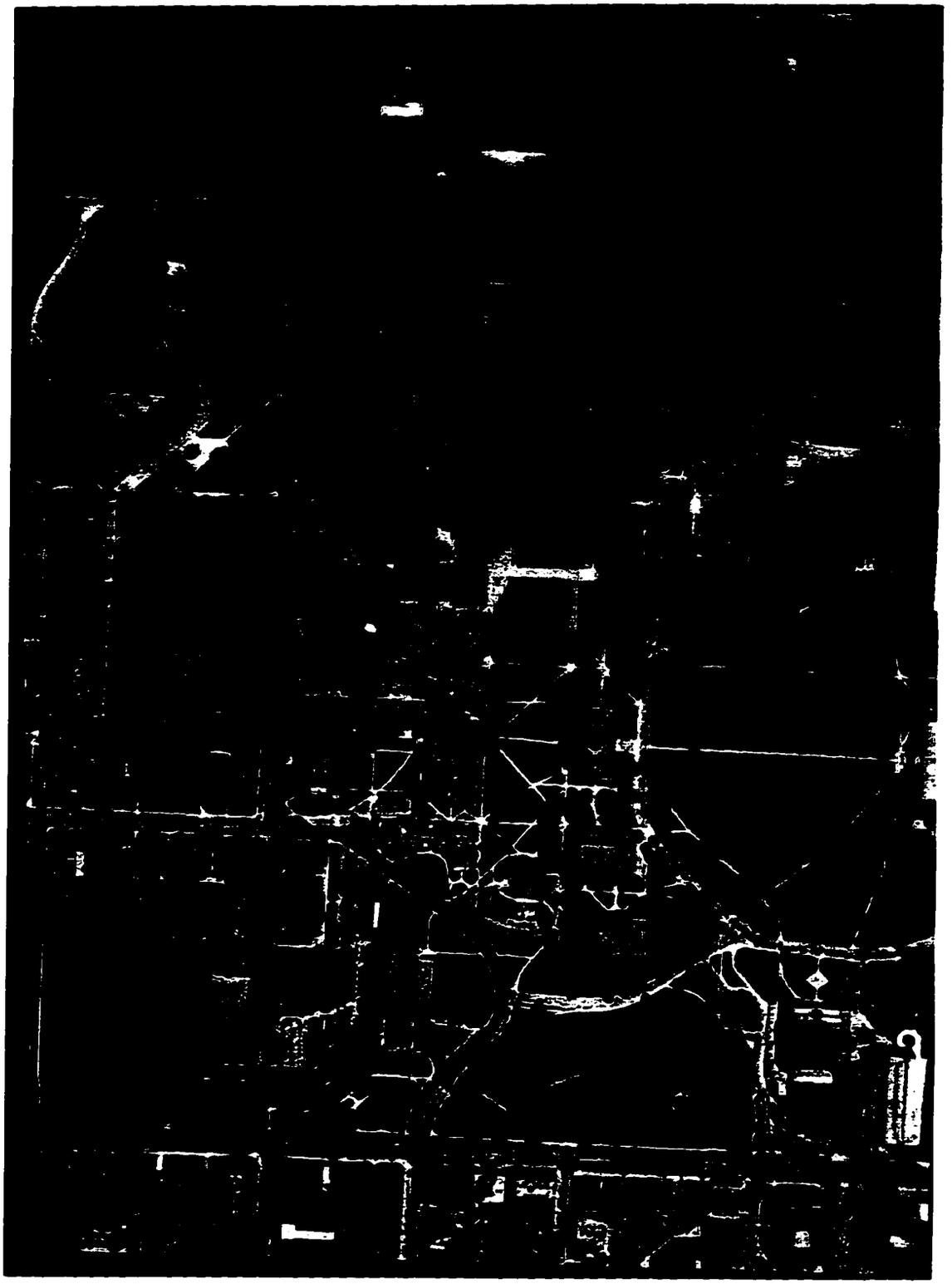


Figure 4.8 New Common Area on Photograph No. 2

4.5.1 Comparison of the Rectification of Single Photograph Using SoftPlotter

In order to check the rectification of each single photograph, the photographs in Figures 4.7 and 4.8 are imported to the SoftPlotter and were rectified again by means of ground control points.

Since the exterior parameters of each photograph used in the above processing were from the SoftPlotter, the rectification in the SoftPlotter for the photographs in Figures 4.7 and 4.8 should give the results shown in the expected column in Table 4.5 if the same ground control points are used again and the rectification discussed in Chapters 3 and 4 are correct. 2.754570 degrees in the table is the angle from the epipolar transformation. The results from SoftPlotter are also listed in Table 4.5

Table 4.5 Comparison between Expected Values and SoftPlotter Results

Name	Photograph No. 1		Photograph No. 2	
	Expected	SoftPlotter Results	Expected	SoftPlotter Results
X_0 (meters)	1487238.563	1487236.119	1487792.068	1487791.914
Y_0 (meters)	1058385.185	1058386.057	1058411.816	1058412.801
Z_0 (meters)	1229.815	1228.750	1229.815	1229.692
ω (degrees)	0	-0.052169	0	-0.084324
ϕ (degrees)	0	-0.161815	0	-0.028110
κ (degrees)	2.754570	2.746991	2.754570	2.735340

The standard error for the locations of both photographs is about 1.283 meters; the standard error for the angles is about 0.057351 degree or 0°03'26.47". The differences probably are caused by the following reasons:

- In the SoftPlotter, the control points for both models are not exactly the same since one vertical control point on the original model does not appear on the compared photographs. The difference in the control point is shown in Table 4.6. The difference could be the main reason. In Table 4.6 the blank cell means that these values are not used as control value. The only difference is that point 163 is used as horizontal control in the original model, but in the check model the point 161 is used as horizontal control instead of the point 163. This difference in control results in the differences of the aerial triangulation for these two cases. If

the flight height above ground is about 750 meters; the angle differences of both ω and ϕ are about 0.057351 degrees, then the estimated distance on the ground is about 1.06 meters.

- All the points on the photographs to be checked are remeasured. These new measurements can be slightly different from those on the original photographs. The approximate formula for calculating the height difference caused by the factor is

$$\Delta H = \frac{Bf}{P_x^2}\Delta p_x$$

Where Δp_x is the difference in measurement; airbase B is the distance between these two exposure stations, which is about 554 meters; P_x is the x parallax; and f is the focal length equal to 152 mm. If the x parallax for the flight height 750 meters above average ground height is used, which is about 111.4667 mm, and the pixel measurement difference is about 0.05 mm, the difference in height is about 0.336 meters. The difference in XY plane is approximately equal to

$$\frac{C_{xr}(H-h_a)}{f}\sqrt{2}$$

which is about 0.349 meters. Where H-h_a is the flight height above the average ground height, which is about 750 meters.

From the above discussion it is seen that these two reasons can account for about $\sqrt{1.06^2+0.339^2} = 1.113$ meters, which is close to 1.283 meters found as the standard error. So we may conclude that the rectification of single photograph in Chapter 3 was correct. Since the method used for calculating the precision is an approximation, another check is presented in the next section.

4.5.2 Comparison of the Rectification of A Single Photograph with the Original Model

Another way to check the correctness of the rectification of a single photograph is to compare the ground coordinates from the stereo pair in Figures 4.7 and 4.8 with these of original model in the SoftPlotter. The check performed in the previous section is on the location and orientation of the photographs, or the exterior orientation parameters. In this section, the check is on the location of the control points in the ground coordinate system. Since the same exterior orientation parameters are used in both models, the ground

coordinates of the same point should be the same for both models if the rectification is correct and the measurements of these ground points are same on both models. The comparison for those control point locations computed after rectification is shown in Table 4.7. From Table 4.7 the standard errors of the differences of the listed control points for X, Y and Z are 0.159 meters, 0.225 meters and 0.209 meters, respectively. From the discussion in the previous section it is known that one pixel measurement difference can bring about 0.34 meters difference in each direction for a point, which is larger than the difference from Table 4.7. So it is reasonable to conclude that the one of the reasons for the difference in the comparison is the measurement difference between these two models. Therefore the rectification of single photograph presented in Chapter 3 agrees with the rectification done by the SoftPlotter.

4.5.3 Check of *y* Parallax for Some Selected Points

The *y* parallax is the difference of *y* coordinates for the same point on each photograph in the stereo pair in epipolar coordinate system. If the epipolar rectification is successful, the *y*

Table 4.6 Comparison for the Difference in Control Points

Point No.	Original Model In SoftPlotter			Compared Model in SoftPlotter		
	X	Y	Z	X	Y	Z
161				1487238.331	1058922.364	
162	1487217.821	1058522.447	291.047	1487217.821	1058522.447	291.047
163	1487223.801	1058124.074				
172	1487494.664	1058486.702	290.610	1487494.664	1058486.702	290.610
173	1487748.207	1058082.750	284.719	1487748.207	1058082.750	284.719

Table 4.7 Comparison with the Original Model

Point No.	From Original Model in SoftPlotter			Computed from the Stereo Pair in Figures 5.7, 5.8		
	X	Y	Z	X	Y	Z
161	1487237.923	1058922.102	292.434	148237.63	1058922.25	292.336
162	1487217.875	1058522.182	291.021	1487217.38	1058522.38	290.903
172	1487494.556	1058486.756	290.659	1487494.25	1058487.00	290.216
1710	1487640.188	1058930.648	288.443	1487640.00	1058931.00	287.972
173	1487748.160	1058082.621	283.528	1487748.13	1058082.75	283.511
1720	1487643.620	1058493.276	290.455	1487643.25	1058493.50	290.442

parallax should be zero for all points. To check this, some points are randomly selected and their y coordinates are shown in Table 4.8.

In Table 4.8 the points beginning with letter "t" are the tie points and the others are control points. The tie points beginning with "t1" are the ground points which can be easily identified, such as the corner of lawns, intersections. The tie points beginning with "t2" are the points on lawns, trees. The tie points beginning with "t3" are the points on the roof of buildings. The y coordinate of each point is not shown in Table 4.8. To obtain the y coordinates of these points, equation $y_e = y_{0e} + J_e C_{xy}$ is used. The difference between the y coordinates of each pixel on all the photographs in the stereo pair is the y parallax.

Table 4.8 Check of the Result of Epipolar Rectification

Point No.	Photograph No. 1		Photograph No. 2		y Parallax (pixels)
	Pixel x	Pixel y	Pixel x	Pixel y	
161	274.87	732.37	238.62	594.37	0.66
162	145.62	2041.62	112.16	1903.26	1.02
172	1045.12	2201.37	1012.87	2064.12	-0.09
1710	1588.12	775.12	1560.12	636.87	0.91
173	1799.87	3556.62	1780.37	3418.62	0.66
1720	1533.62	2203.37	1500.87	2065.87	0.16
t11	836.25	698.25	800.67	560.41	0.50
t12	114.75	2034.75	81.25	1896.75	0.66
t13	362.75	3538.75	336.93	3400.67	0.74
t14	1174.75	3480.25	1157.44	3342.74	0.17
t15	1267.75	2021.25	1234.83	1882.79	1.12
t16	2412.25	1044.25	2391.02	905.04	1.87
t17	2479.75	2969.75	2458.90	2831.48	0.93
t31	682.25	1066.75	605.75	927.75	1.66
t21	137.00	1094.00	97.92	954.54	2.12
t32	362.00	2247.00	300.00	2110.00	-0.34
t22	482.75	2180.25	446.60	2042.01	0.90
t23	413.00	2805.00	377.50	2665.19	2.47
t33	1168.00	1979.00	1033.00	1843.00	-1.34
t34	1719.00	2623.00	1648.49	2486.05	-0.39
t36	2042.00	1150.00	1991.00	1012.00	0.66
t24	2239.00	1853.00	2207.58	1713.81	1.85
t25	2371.00	2299.00	2349.25	2155.10	6.56
t26	2318.25	2859.75	2295.74	2720.31	2.10

The standard error for all the points in Table 4.8 is 1.318 pixels. In this case, the search range in *y* axis can be limited within 2 pixels in the image matching. The reasons for the *y* parallax could be that the relative orientation of the model is not perfect and the measurements are not perfect.

Since these points are all over the whole area, this *y* parallax standard error is acceptable.

4.6 Direct Transformation from Interior Orientation to Stereo Rectification

The resampling methods are used in each step of the image processing on the aerial photographs. This is not good for the image matching because resampling changes the pixel gray value of each location on the photograph by interpolation, and it is the pixel gray values that are used in the image matching. Therefore it would be better to do image processing in one step, directly from the raw photographs to the stereo rectified photographs, to make the affection from resampling to the least.

Another advantage of direct transformation is that it can save much time that would be spent in the resampling process for each step. Due to the huge image size, the resampling process for each step could take a lot of time.

Since the final processing is stereo rectification, more than two images are involved in the one step image processing; at least two images are needed as the input parameters of the processing. The interior and exterior orientation parameters for each photograph in the stereo pair must be known before the direct transformation.

4.6.1 Direct Transformation Formula

Matrix operation is used in the formula to perform the direct transformation from the interior and exterior orientation to the epipolar rectification.

$$\begin{bmatrix} x_e \\ y_e \\ z_e \end{bmatrix} = \frac{1}{k}\begin{bmatrix} x'_e \\ y'_e \\ z'_e \end{bmatrix} = \frac{S}{k}\begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix} R^{-1}\left(\begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} I \\ J \\ -f \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ 0 \end{bmatrix}\right) \qquad 4.16$$

Where x_e, y_e, z_e are the epipolar coordinates; S is the epipolar rotation matrix from equation 4.12; R is the rectification rotation matrix from equation 3.4; k is the scale between the model and the ground; s is the equalizing scale from equation 4.6, for the reference

photograph it equals 1; a, b, c, d, x_0, y_0 are the transformation parameters from the control points for the interior orientation; I, J are the pixel coordinates before the interior orientation. Then the values of x_e, y_e are:

$$x_e = \frac{x_e'}{z_e'} z_e$$
$$y_e = \frac{y_e'}{z_e'} z_e \qquad 4.17$$

Notice that the z_e is the value for z coordinates of all image pixels on the photograph in the epipolar coordinate system. Normally it is equal to $-f$ for the same scale transformation. The equations 4.16 and 4.17 can be applied to each photograph in the stereo pair with different transformation parameters.

4.6.2 Determination of the Pixel Size

Equations 4.16 and 4.17 only gives the coordinate relationship between the raw pixel coordinates and the epipolar coordinates. In order to represent the digital photograph, the constants of the pixel coordinates associated with the epipolar coordinates have to be available, which means the pixel size, location of the initial pixel in the epipolar coordinate system and the size of the photograph in the epipolar coordinate system have to be known before the direct transformation.

Equation 2.6 determines the pixel size after interior orientation. Since there are more than one photograph are used in the direct transformation process, assuming C_{xr}, C_{yr} are the pixel size of the reference photograph and C_x', C_y' are the pixel size to be changed. After the correction of the pixel size, all the photographs should have the same pixel sizes of C_{xr}, C_{yr}.

If the absolute rotation angle around z axis in the direct transformation is less than 45° or larger than 135° and less then 180°, the pixel size of the epipolar images will remain the same. To determine this, assuming $I = 1$, $J = 0$, $x_0 = y_0 = 0$, equation 4.16 becomes

$$\begin{bmatrix} x_e \\ y_e \\ z_e \end{bmatrix} = \frac{S}{k} \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix} R^{-1} \begin{bmatrix} a \\ c \\ -f \end{bmatrix} \qquad 4.18$$

Once the values of x_e and y_e are available, the pixel sizes in the epipolar coordinate system are:

$$\begin{aligned} &if \quad |x_e| > |y_e| \\ &\quad C_{xe} = C_{xr} \\ &\quad C_{ye} = C_{yr} \\ &else \\ &\quad C_{xe} = |C_{yr}| \\ &\quad C_{ye} = C_{xr} \frac{C_{yr}}{|C_{yr}|} \end{aligned} \qquad 4.19$$

Therefore the final pixel size can be determined. Notice that the C_{xr}, C_{yr} are the pixel size of the reference image in the stereo pair which needs to be determined before the direct transformation.

4.6.3 Determination of the Initial Location and the New Image Size

To determine the coordinates of the initial locations of the photograph in the stereo pair in the epipolar coordinate system, one needs to determine the limits of each photograph in the epipolar coordinate system. Assume the sizes of one photograph in the stereo pair before the direct transformation are ***size_x***, ***size_y***, then the limits of x_{min}, x_{max}, y_{min}, y_{max} in the epipolar coordinates are available by substituting 0, ***size_x*** - 1 and 0, ***size_y*** - 1 in equation 4.16, then the initial location x_{0e}, y_{0e} for this photograph are:

$$\begin{aligned} &x_{0e} = x_{\min} \\ &y_{0e} = y_{\min} \; if \quad C_{ye} > 0 \\ &y_{0e} = y_{\max} \; if \quad C_{ye} < 0 \end{aligned} \qquad 4.20$$

and the new photograph size are:

$$\begin{aligned} size\_x_e &= \frac{x_{\max} - x_{\min}}{C_{xe}} \\ size\_y_e &= \frac{y_{\max} - y_{\min}}{|C_{ye}|} \end{aligned} \qquad 4.21$$

Although the pixel size for all the photographs in the stereo pair must be the same, the initial location and the image size of them are different.

4.6.4 Resampling Equation for the Direct Transformation

Assume

$$D=\begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = S\begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix} R^{-1} \qquad 4.22$$

for the process of single photograph rectification, equalizing scale and epipolar rectification, then equation 4.16 becomes

$$\begin{bmatrix} x_e \\ y_e \\ z_e \end{bmatrix} = \frac{D}{k}\left(\begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} I \\ J \\ -f \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ 0 \end{bmatrix}\right) \qquad 4.23$$

If

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = kD^{-1}\begin{bmatrix} x_{0e} + C_{xe}I_e \\ y_{0e} + C_{ye}J_e \\ z_e \end{bmatrix} for$$

$$I_e = 0,\cdots,size\_x_e \qquad 4.24$$

$$J_e = 0,\cdots,size\_y_e$$

where z_e normally equals $-f$. Then equation 4.23 becomes:

$$\begin{bmatrix} \frac{x'}{z'} \\ \frac{y'}{z'} \end{bmatrix} = \begin{bmatrix} \frac{a}{-f} & \frac{b}{-f} \\ \frac{c}{-f} & \frac{d}{-f} \end{bmatrix}\begin{bmatrix} I \\ J \end{bmatrix} + \begin{bmatrix} \frac{x_0}{-f} \\ \frac{y_0}{-f} \end{bmatrix}$$

It can be proved that $z' = -f$ from equations 4.23 and 4.24, therefore the final equation for the resampling is:

$$\begin{bmatrix} I \\ J \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}\begin{bmatrix} x' - x_0 \\ y' - y_0 \end{bmatrix} \qquad 4.25$$

The resampling formula for the direct transformation consists of two equations, equations 4.24 and 4.25. Equation 4.24 gives the values of x', y' for the equation 4.25. The comparison between the direct transformation and the results from the multi-procedures discussed in Chapters 2 and 3 are shown in Table 4.9.

Table 4.9 Comparison between Direct Transformation and the Multi-Procedures

Pt No.	Direct Transformation				Multi-Procedures			
	Photograph No. 1		Photograph No. 2		Photograph No. 1		Photograph No. 2	
	x (mm)	y (mm)	x(mm)	y (mm)	x (mm)	y (mm)	x (mm)	y (mm)
161	4.077	87.254	-86.034	87.232	4.052	87.248	-86.064	87.281
162	-2.304	22.390	-92.278	22.392	-2.356	22.413	-92.334	22.463
172	42.292	14.505	-47.654	14.465	42.241	14.502	-47.677	14.497
1710	69.215	85.155	-20.530	85.109	69.163	85.131	-20.544	85.176
173	79.665	-52.616	-9.610	-52.632	79.661	-52.611	-9.624	-52.579
1720	66.505	14.394	-23.428	14.363	66.461	14.403	-23.482	14.411

The standard error of the difference between these two sets of results is about 0.041 mm or 0.83 pixel. Again there is a re-measurement problem here which could cause error around one pixel or less. So the comparison shown in Table 4.9 the agreement on geometric transformation between these two methods. Table 4.10 shows the radiometric comparison between these two methods.

Table 4.10 Comparison between the Radiometric Transformations

Point No.	Originals		Direct Transformation		Multi-Procedures	
	Photo 1	Photo 2	Photo 1	Photo 2	Photo 1	Photo 2
	Pixel Radiometric Value					
161	110	88	110	82	146	172
162	171	157	169	155	86	102
172	159	131	149	132	106	127
1710	163	157	151	148	104	106
173	131	123	133	122	123	132
1720	123	83	119	82	132	173

Although the photographs after the multi-procedure processing virtually are the same as the originals on the contrast for the above points, their pixel gray values have been changed much more than the change resulted from the direct transformation. So the direct transformation preserves the image better for image matching.

4.7 Conclusions

Stereo rectification and direct transformation are discussed and applied in this chapter. Equalizing pixel size and photograph scale in the stereo rectification makes pixels the distance

units that can be used in the image matching. An error from the stereo rectification is discussed and its correction is also given.

In the direct transformation, one resampling brings the digital photograph from the raw image to the image after stereo rectification which is ready for image matching. This method decreases resampling to one time, therefore it has the least effect on the radiometric characteristics, which is good for image matching; and it also need much less resampling so that much time can be saved.

The results shown above indicate that the direct transformation is consistent with the multi-procedures on the geometric transformation, but the radiometric transformation is different between them. The direct transformation shows much better results on preserving the original image information.

The direct transformation formulas are not simply the combination of all the formulas in the mutli-procedure since more than one photographs are involved in the stereo rectification. It has big advantage over the multi-procedure methods.

CHAPTER 5 INITIALIZATION OF IMAGE MATCHING

Before performing image matching, certain information about the stereo pair should be given. Some of the information include the ground height variance, the average ground height, the ground spacing of the output DEMs, matching window size.

5.1 Determination of the Ground Coordinates

After image matching, the ground height for each pixel on the reference photograph can be obtained. But the x and y coordinates are still in the image coordinate system, such as epipolar coordinate system. We need to compute the X and Y coordinates of these pixels and convert them to the ground coordinate system.

5.1.1 Ground Height from the x Parallax

Figure 5.1 shows the geometry in a stereo pair of photographs. Where f', f are the principal distances for these two exposures; H', H are the flight heights above Mean Sea Level (MSL); B is the airbase or the distance between these two exposure stations; A is a ground point; a', a are its image points on the left-hand and right-hand photographs, respectively; x', x are the x coordinates in the epipolar coordinate systems on the left-hand and right-hand photographs, respectively. Assuming the ground locations of the left-hand and the right-hand photographs are (X', Y', Z') and (X, Y, Z), respectively, according to Figure 5.1 the x parallax of the point A is determined as (x' - x) if they have the same flight height and principal distance. Since after epipolar rectification the scale of all the photographs except the reference photograph in the stereo pair may be changed by the scale equalizing factor s, the x parallax may be slightly different. Using the left photograph as the reference photograph, the ground height of the point A is given by:

$$h_A = \frac{(fZ'x' - \frac{x}{s}Zf') - Bff}{fx' - \frac{x}{s}f'} \qquad 5.1$$

for the case shown in Figure 5.1. Where s is the scale equalizing factor discussed in the Chapter 4. Notice that it is important to include the s in calculating the heights for the correct ground height. In the equation 5.1, x divided by s is the correction applied to remove the error introduced in the scale equalizing process. Z in the above equation is the real flight height above MSL for the right-hand camera, not the height after equalizing scale process.

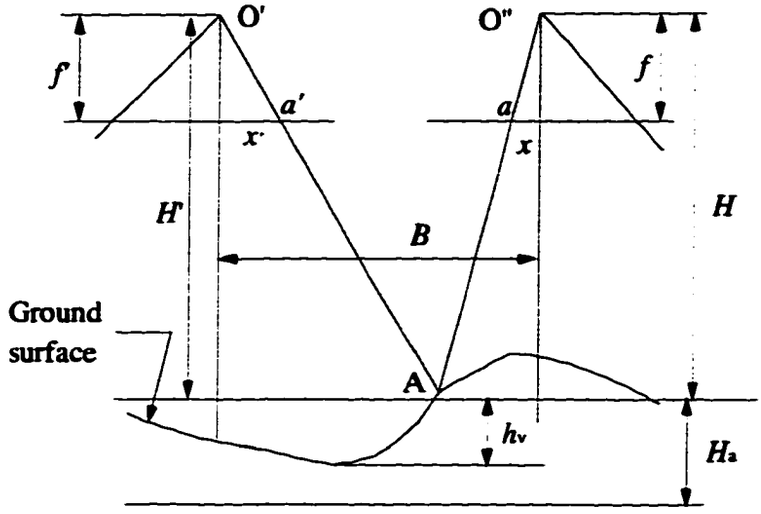


Figure 5.1 Average Ground Height and Its Variance

5.1.2 Calculation of the Ground X and Y Coordinates from Matched Ground Height

The height of a point in the stereo pair can be determined with equation 5.1 or equation 5.2 once the image matching is done and the x parallax is available. But this does not change the x, y location of the pixels on the reference image. In order to get the ground coordinates of each pixel in the stereo pair, the rectified coordinates must be restored from the epipolar coordinate system. Assuming that the left photograph in Figure 5.1 is the reference photograph in the stereo pair, and the pixel coordinates of the point a' is I', J'; the initial location of the left photograph is x'_0, y'_0; the pixel sizes are $C_{x'}$, $C_{y'}$; the angle rotation in the epipolar rectification process is θ; and the ground coordinates of the point A is X_A, Y_A, the ground X_A and Y_A can be expressed as:

$$\begin{bmatrix} X_A \\ Y_A \end{bmatrix} = \frac{Z_r - h_A}{f'} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} I'C_{x'} + x'_0 \\ J'C_{y'} + y'_0 \end{bmatrix} + \begin{bmatrix} X_r \\ Y_r \end{bmatrix} \qquad 5.2$$

Where X_r, Y_r, Z_r are the ground location of the reference photograph. Then all the ground coordinates of each matched pixel on the reference image are available.

5.2 Initialization of Image Matching

Before the image matching can get started, some parameters of the stereo model and the image searching must be determined. Those parameters include the average ground height, the maximum variance of the ground around the average ground height, the x parallax on that average ground height, the searching range or pull-in distance for given maximum variance of the ground surface, the ground spacing of the output DEM, the maximum slope expected in the output DEM, searching window sizes in terms of number of row and column, initialization of the LAN header of each image file, number of bits per pixel of output DEM and so on.

5.2.1 Average Ground Height and the Ground Variance

The average ground height H_a above the MSL determines the average scale of each photograph in the stereo pair; and the maximum ground variance determines the largest scale variance of the photograph. As shown in Figure 5.1, H_a is the average ground height; h_v is the maximum deviation of the ground surface from the average ground height; H is the flying height above the average height.

If the flight height above MSL is Z_0, then

$$H = Z_0 - H_a$$

and

$$V = \frac{h_v}{Z_0 - H_a} \qquad 5.3$$

is defined as the variance of the model. In this research the flight height above the MSL is about 1230 meters, the average ground height above the MSL is about 290 meters; if $V = 0.1$, then the maximum ground variance from the average ground height is about 94 meters; and the total variance for the ground is 188 meters. The value of V is directly related to the searching range in the image matching. The larger the V, the larger the searching range. The value of V depends on the type of ground features. For flat area, a small V may be good

enough since the height of the ground surface is closely around the average height. For a hilly areas or large scale areas like cities, V needs to be larger.

5.2.2 Estimation of the Initial Matching Positions

For the given pixel position I' on the reference photograph and the flight height above the average ground ($Z' - h_a$), the initial matching pixel position I on the photograph to be matched can be calculated as:

$$I = I' + \frac{x_0' - x_0 - \frac{Bf'}{Z' - h_a}}{C_x} = I' + t \qquad 5.4$$

Where x'_0, x_0 are the initial x coordinates of reference and matched photographs, respectively; C_x is the pixel size in x direction; t here is used as a constant for shifting the position on the reference photograph to obtain the initial estimated position on the matched photograph.

5.2.3 Estimate of the Searching Range

If the value of V and the flight height above the average ground height is given, from equation 5.4 the searching range on the photographs to be matched can be determined by

$$\Delta I = \frac{BfV}{C_x(Z' - h_a)} \qquad 5.5$$

in number of pixels. For this case, B = 554 meters, f' = 152.473 mm; (Z' - h_a) = 940 meters; C_x = 0.04958 mm; if V = 0.05, then the value of ΔI = 90.6 pixels on the row, which means that the matching position is supposed be within 90.6 pixels from the initial estimated position. It is not always necessary to search such a large range. For a matched pixel, the next matching pixel should be very close since the surface of the ground is continuous, except for the walls of buildings.

The searching range in the y or the columns depends on the accuracy of the relative orientation of the model. the less the y parallax is, the smaller the y search range is needed. It should be less than 5 pixels in a normal situation.

5.2.4 Ground Spacing

The ground spacing determines the spacing of the output DEM raster. Assume the ground spacing in X and Y is C_X, C_Y, the numbers of pixels within one ground spacing are:

$$n_X = \frac{C_X f}{C_x H}$$
$$n_Y = \frac{C_Y f}{C_y H} \qquad 5.6$$

Where H is the flight height above the average ground height; C_x, C_y are the pixel sizes. If n_X, n_Y are larger than 1, only every n_X, n_Y pixels needs to be matched. Otherwise all the pixels must be matched.

5.2.5 Maximum of the Ground Slope

For some types of features like buildings, the slope angle for the walls is almost 90°. The large slope on the ground can cause the discontinuity or dissimilarities on the photographs in the stereo pair. On the other hand, the larger the slope angle is, the larger the searching range is needed to match the neighbor pixels. So it is reasonable to set the maximum slope in the image matching process. From equation 5.5, if the maximum ground slope is α, for the given ground spacing C_X, the ground height difference is $\tan(\alpha)C_X$. So the number of pixels caused by this height difference is:

$$\Delta I = \frac{Bf C_X \tan(\alpha)}{C_x (Z' - h_a)^2}$$

Substituting equation 5.6 into the above equation:

$$\Delta I = \frac{B n_X \tan(\alpha)}{(Z' - h_a)} \qquad 5.7$$

If the number of pixels for a given ground spacing $n_X = 2$, slope angle $\alpha = 85°$ and $\tan(85°) = 11.43$, the searching range is about 13.47 pixels that is much less than the searching range shown in section 5.2.3.

The maximum slope can be defined according to the type of ground features and any other particular requirement.

5.2.6 Window Size

A window is defined before the image matching process as a mask. All the pixels in the window are used in the process of matching. Generally speaking, the larger the window, the more time is consumed in the process and the better is the matching. Therefore, from the point of view of processing speed, the smaller the window, the better; but for matching precision, the larger window, the better matching results. Normally, the larger the window is, the more the similarity and dissimilarity will be from the ground features. If the increase of dissimilarity is more than the increase of similarity then increasing the window size, the matching precision will decrease for larger window. The information in the window may be not enough for a good matching if the window is too small; the window size should be big enough to have enough information in the window so that the image matching can be successfully done. Since all the pixels except the matched pixel are geometrically distorted because of the ground relief, and some other noises that might appear in the window, smaller window are better. So it is necessary to balance between the above factors and determine the optimal window size according to the situation.

Although the number of rows and columns of a window can be an even number, the odd number is preferred for simplification. This means only odd number like 3, 5 should be used as the size of a window. The only difference between these two types of window is the calculation of the matched position. The shape of a window can be square or non-square. Both square and non-square windows will be discussed in Chapter 8.

5.2.7 Number of Bits Per Pixel in Output DEM

The number of bits per pixel in the output DEM determines the maximum height variance and precision of the output DEM. For example, a 8 bits number per pixel can represent a maximum number of 255 which can represent a maximum variance of 256 meters if precision of the ground is 1 meter. In order to reduce the storage space for the output DEM, integers are used. In practice, a precision less than 1 meter may be needed. To solve this problem, assume there is a scale factor ***Scale*** between the real ground height and the pixel gray value in DEM as

$$Z = \frac{g(I,J)}{Scale} \qquad 5.8$$

Where Z is the real ground height; $g(I, J)$ is the pixel gray value. Therefore the integer values of pixels can be used to represent floating numbers that have precision of less than 1 meter. For example, the real height is 300.03 M if the pixel gray value is 30003 and the ***Scale*** is 100. Therefore the precision actually is the inverse of the value of the ***Scale***, in this example 0.01M. In order to have sufficient bits to accommodate the largest possible height variance on the ground, 16 bits or 24 bits per pixel can be used to represent maximum of 65535 or 16777215 depending on the height variance and the required precision. To extend the expressing range of a 16 bits number, equation 5.9 is used. In equation 5.9 the difference between the original height value and the minimum height in the stereo pair, ***Min***, is used to replace the original height value. Since $g'(I, J) <= g(I, J)$ is always correct, the range of height is larger than it was.

$$Z = \frac{g'(I,J) + Min}{Scale} \qquad 5.9$$
$$g'(I,J) = g(I,J) - Min$$

5.2.8 Checking on the Overlapping Area of Photographs in A Stereo Pair

The image matching can be done successfully only if all the photographs in a stereo pair have the same overlapping area. Otherwise the image matching will definitely fail or give the wrong results. In order to check this, the flowchart shown in Figure 5.2 is used to check if two photographs in a stereo pair have common overlapping area on the average ground height h_a. The basic idea here in Figure 5.2 is to check to see if the extreme area on one photograph can be transferred to another.

5.2.9 Selection of Initial Points

In order to obtain some information about the stereo model to be matched, a certain number of initial points can be used. To select the initial points, evenly distributed points on the stereo pair are used. The simplest way to locate these initial points is to put these points on the pixels. Assuming the number of initial points is m; image size is ***size_x***, ***size_y***, let the m initial points be arranged as m_x by m_y matrix on the reference image in the stereo pair, the values of m_x and m_y are determined by:

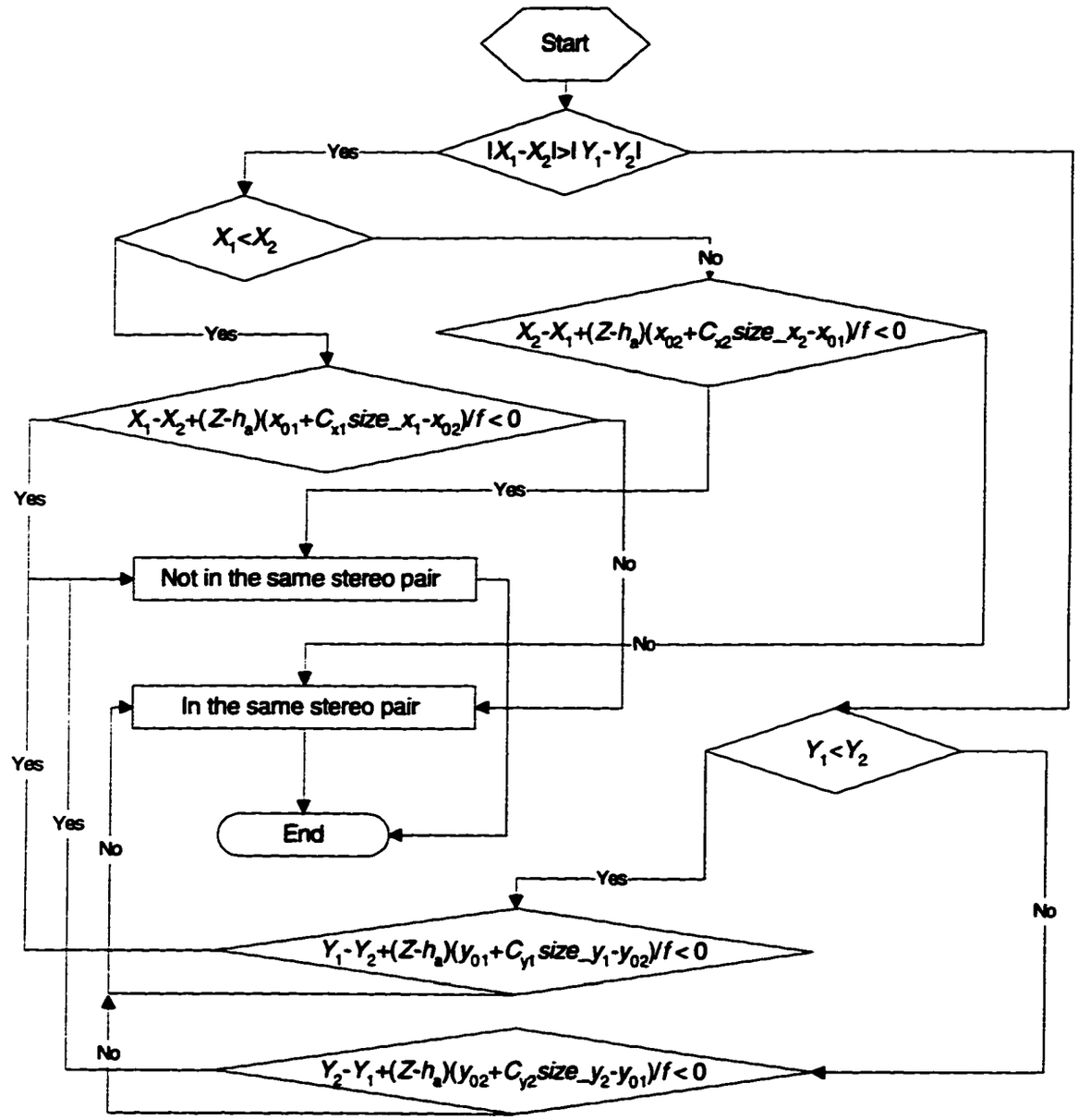


Figure 5.2 Flowchart for Checking on Two Photographs in the Stereo Pair

$$m_x = \frac{size\_x}{size\_x + size\_y} m$$
$$m_y = \frac{size\_y}{size\_x + size\_y} m \qquad 5.10$$

and the pixel locations of these m points are:

$$I_i = \frac{size\_x - 1}{m_x + 1} i$$
$$J_j = \frac{size\_y - 1}{m_y + 1} j \qquad 5.11$$
$$for \quad i = 1, \cdots, m_x; j = 1, \cdots, m_y$$

Image matching can be performed first on these initial points to find out some characteristics of the images in the stereo pair.

5.2.10 Using the Initial Points to Determine Some Threshold Values

The threshold values from the initial points will be used to determine if the image matching results are acceptable.

Using the initial points from the previous section, the coefficients of the cross correlation can be found for those points in the cross correlation image matching; or the least differences in the least difference image matching are available. Assuming the coefficients of cross correlation for these initial points are r_{ci} for i = 1 to m , the condition can be set as:

$$r_c \geq r_a - \delta_r \qquad 5.12$$

Where r_a is the average of these coefficients; δ_r is the standard error of these coefficients. It is known that the larger the coefficient r is, the more probable it is that the matching is correct. For any matching with a coefficient of matching that meets equation 5.12, the matching is considered correct at a confidence of 84.13% assuming the normal distribution. Here difference confidence level can be set for different matching qualities.

The story is the same for the least difference image matching. Assuming the differences for those initial points are d_i; d_a is the average difference; δ_d is the standard error. A condition similar to equation 5.12 can be written as

$$d \leq d_a + \sigma_d \qquad 5.13$$

Differing from equation 5.12, the smaller the difference is, the higher probability the matching is correct.

After the results from the equation 5.12 or 5.13 for the initial points are available, it can also be used to check the points so that those initial points that do not meet the condition can be removed. Sometimes, the matching could be wrong even when the matching shows very good results with very large coefficients or small differences. To detect this, multi-method is used which will be discussed in Chapter 7.

5.2.11 The Check Points Used in the Study

The check points are those listed in Table 4.8. Here the photo pixel coordinates and the ground coordinates of these points on photographs in Figures 4.7 and 4.8 are listed in Table 5.1.

Table 5.1 Pixel and Ground Coordinates of Check Points

Point No.	Photograph No. 1		Photograph No. 2		Ground Coordinates		
	Pixel x	Pixel y	Pixel x	Pixel y	*X*(meters)	*Y*(meters)	*Z*(meters)
161	274.87	732.37	238.62	594.37	1487237.63	1058922.25	292.336
162	145.62	2041.62	112.16	1903.26	1487217.38	1058522.38	290.903
172	1045.12	2201.37	1012.87	2064.12	1487494.25	1058487.00	290.216
1710	1588.12	775.12	1560.12	636.87	1487640.00	1058931.00	287.972
173	1799.87	3556.62	1780.37	3418.62	1487748.13	1058082.75	293.511
1720	1533.62	2203.37	1500.87	2065.87	1487643.25	1058493.50	290.442
t11	836.25	698.25	800.67	560.41	1487408.13	1058941.00	291.952
t12	114.75	2034.75	81.25	1896.75	1487207.88	1058524.00	290.925
t13	362.75	3538.75	336.93	3400.67	1487305.88	1058068.13	286.917
t14	1174.75	3480.25	1157.44	3342.74	1487555.25	1058096.63	282.400
t15	1267.75	2021.25	1234.83	1882.79	1487559.50	1058545.13	290.548
t16	2412.25	1044.25	2391.02	905.04	1487898.63	1058862.75	284.379
t17	2479.75	2969.75	2458.90	2831.48	1487947.88	1058273.13	284.149
t31	682.25	1066.75	605.75	927.75	1487363.88	1058816.13	312.629
t21	137.00	1094.00	97.92	954.54	1487201.00	1058809.50	293.803
t32	362.00	2247.00	300.00	2110.00	1487285.75	1058461.75	305.433
t22	482.75	2180.25	446.60	2042.01	1487322.13	1058484.88	292.270
t23	413.00	2805.00	377.50	2665.19	1487310.13	1058293.88	291.940
t33	1168.00	1979.00	1033.00	1843.00	1487513.00	1058547.38	340.604
t34	1719.00	2623.00	1648.49	2486.05	1487696.50	1058368.63	309.588
t36	2042.00	1150.00	1991.00	1012.00	1487777.50	1058817.63	299.766
t24	2239.00	1853.00	2207.58	1713.81	1487853.88	1058610.75	289.705
t25	2371.00	2299.00	2349.25	2155.10	1487904.25	1058477.25	284.654
t26	2318.25	2859.75	2295.74	2720.31	1487896.00	1058304.50	285.056

From the previous chapters it is know that the coordinates of these check points can be considered correct. So they will be used in this study as the check for the image matching results. As mentioned before, all the points beginning with "t1" are the highly contrasted ground points; all the points beginning with "t2" are the ground points in the area of grass and trees; and all the points beginning with "t3" are the points on the top of buildings. The highest and lowest points in the above table are 't33' and 't14' with the heights of 340.604 meters and 282.400 meters, respectively. Therefore the height variance of the area of interest is about 58 meters.

5.2.12 The Common Area to Be Used to Create DEM

A smaller test area is selected from the rectified stereo pair to create the DEM. The selected area is shown in Figures 5.3 and 5.4. In this area almost all the types of ground features in the flight project, such as buildings, roads, grass and trees, are available. So the image matching on the selected area should be typical for the whole model area.

5.3 Conclusions

Image matching is a process to find the best-matched point on the photograph to be matched pixel by pixel in the searching range. Some initial parameters, such as the average ground height, the ground variance, are set in this chapter so that the image matching can be performed. A number of pixel points are used to determine the characteristics of the images in the stereo pair. The threshold values used to check the matching results can be found from these initial points. For this research, only 24 given points are used as the initial points. Both the pixel and photo coordinates of the 24 check points are given. Two images in a common area that can be used to create DEM are generated.

Figure 5.3 Selected Area from Photograph No. 1

Figure 5.4 Selected Area from Photograph No. 2

CHAPTER 6 BASIC IMAGE MATCHING ALGORITHMS AND SEARCHING STRATEGIES USED

Only three types of area based image matching methods, cross correlation, least difference and least square image matching, are used in this study. The methods to be discussed in the following chapters can be applied to any other area based image matching algorithms. Two types of searching strategies, whole range and best-track, are discussed. The differences between these image matching algorithms and searching strategies will be shown in the matching results. The 24 check points given in the previous chapter are used as the matching control points for checking the matching results. The height of each point after matching will be compared to the correct height shown in Table 5.1.

6.1 Cross Correlation Image Matching (CCIM)

6.1.1 Normalized CCIM Equation

The equation of cross correlation used here is the normalized cross equation shown in Chapter 1. The only difference is that the square of the coefficient is used in order to reduce the amount of computation, as shown in equation 6.1.

$$R^2(i,j) = \frac{\sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_1(x,y) \bullet g_2(x-i,y-j)}{\left[\sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_1^2(x,y) \bullet \sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_2^2(x-i,y-j)\right]} \qquad 6.1$$

Where $g_1(x, y)$, $g_2(x-i, y-j)$ are the pixel gray values on the stereo pair; i, j are the shift of rows and columns on the image to be matched; (y_1-y_0+1) and (x_1-x_0+1) define the window size. The limits of i, j on the matched image is defined by the search range. The coefficient can not be larger than 1. In order to find out the position i, j with the maximum coefficient value within the searching range, all the calculated coefficients are compared. More information on CCIM is given in Chapter 1.

6.1.2 Determination of Location i_{max} for the Maximum Coefficient

The i, j from equation 6.1 must be integers if no interpolation is applied. Therefore the precision of heights from the integer matching results are confined by at least one pixel. In order to improve the precision of the heights, interpolation is used. Assuming the correlation coefficients of the pixel position with the maximum coefficient and these of two surrounding pixels on the same row are found. The matched location should be within this range of three pixels. In order to find the matched position, parabolic curve is used to do interpolation from these three pixels. The formula for the i_{max} is given by:

$$i_{\max} = \frac{3r(i-1) - 4r(i) + r(j+1)}{2\big(r(i-1) - 2r(i) + r(i+1)\big)} \qquad 6.2$$

Where i-1, i, i+1 and $r(i\text{-}1)$, $r(i)$, $r(i\text{+}1)$ are the pixel locations and the correlation coefficients of the position with maximum coefficient and the two surrounding pixels from equation 6.1; i_{max} is the interpolated location.

6.1.3 Determination of Ground Height from i, j

Once the i, j have be determined, the height of a matched pixel can be calculated by using equation 5.1. The epipolar photo coordinates for the pixels on the reference image and on the matched image are given by:

$$\begin{aligned} x_1 &= x_{10} + C_x i_1 \\ x_2 &= x_{20} + C_x i_{\max} \end{aligned} \qquad 6.3$$

Where x_{10}, x_{20} are the initial x coordinates of the photographs; C_x is the pixel size in x direction; i_1 is the x location of the pixel on the reference image.

Other information such as focal lengths, the airbase, flight, and the epipolar rectification scale of each photograph are available from the information about the stereo pair.

6.2 Least Difference (or Distance) Image Matching (LDIM)

As discussed in Chapter 1, two types of algorithms can be used in LDIM. The normalized equations used in the computer programs are shown in equations 6.4 and 6.5 as:

$$Difference(i,j) = \sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1}\left| g_1(x,y) - \frac{g_2(x-i,y-j)\cdot A_1}{\sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_2(x-i,y-j)} \right| \qquad 6.4$$

$$A_1 = \sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_1(x,y)$$

and

$$Difference(i,j) = \sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1}\left(g_1(x,y) - \frac{g_2(x-i,y-j)\cdot A_1}{\sum_{x=x_0}^{x=x_1}\sum_{y=y_0}^{y=y_1} g_2(x-i,y-j)} \right)^2 \qquad 6.5$$

All the symbols are the same as those used in equation 6.1 except A_1 which is the sum of the pixel gray values for the pixels within the window around the pixel $(\frac{x_0+x_1}{2}, \frac{y_0+y_1}{2})$ on the reference image, is used to normalize the equations. The average pixel gray value on each window is supposed to be used in the normalized equations to calculate the difference or distance. Since the value of A_1 is a constant for each particular pixel on the reference image, it will be easier and faster to use the above format in the computation and will keep the same scale in the comparison between values from different windows in the matching process. Also the square of the distance in the equation is used for less computation time.

Similar to the process in section 6.1, in order to determine the ground height of each matched point, equations 6.2 and 6.3 are used to interpolate the better matching position and determine the parameters needed by equation 5.1 to calculate the ground height of the matched point.

6.3 Least Square Image Matching (LSIM)

The LSIM has been discussed in Chapter 1. The equations used in the computer programs and the problem solving procedures used are explained in this section.

6.3.1 LSIM Equations Used

Assuming pixel location of x_1, y_1 on the reference image is conjugated with a location of x_2, y_2 on the image to be matched and the transformation relationship between them is:

$$\begin{aligned} x_2 &= a_0 + a_1 x_1 + a_2 y_1 \\ y_2 &= b_0 + b_1 x_1 + b_2 y_1 \end{aligned} \qquad 6.6$$

and the pixel gray values between them should have the relationship:

$$g_1(x_1, y_1) = h_0 g_2(a_0 + a_1 x_1 + a_1 y_1, b_0 + b_1 x_1 + b_2 y_1) + h_1 + n_2 - n_1 \qquad 6.7$$

Where g_1, g_2 are the pixel gray values on the reference and matched images, respectively; a_0, a_1, a_2, b_0, b_1, b_2, h_0, h_1 are the geometric and radiometric transformation parameters between these two images; n_1, n_2 are the noise portions of the pixel gray values on these two images. Using Taylor's series, for the estimated initial position (x_2^0, y_2^0) on the matched image, the equation 6.7 can be written as:

$$\begin{aligned} g_1(x_1, y_1) = {} & (h_0 g_2(x_2^0, y_2^0) + h_1) + g_2(x_2^0, y_2^0) dh_0 + h_0 \dot{g}_{2x}(x_2^0, y_2^0) da_0 + h_0 x_1 \dot{g}_{2x}(x_2^0, y_2^0) da_1 \\ & + h_0 y_1 \dot{g}_{2x}(x_2^0, y_2^0) da_2 + h_0 \dot{g}_{2y}(x_2^0, y_2^0) db_0 + h_0 x_1 \dot{g}_{2y}(x_2^0, y_2^0) db_1 \\ & + h_0 y_1 \dot{g}_{2y}(x_2^0, y_2^0) db_2 + dh_1 \end{aligned}$$

if all the small items, including n_2-n_1, are neglected. So basically the least square equation for a pixel on the reference image in a given window can be expressed as equation 8. Where $h_0^0, h_1^0, a_0^0, a_1^0, a_2^0, b_0^0, b_1^0, b_2^0$ and $h_0, h_1, a_0, a_1, a_2, b_0, b_1, b_2$ are the eight transformation parameters before and after each iteration respectively; $dh_0, dh_1, da_0, da_1, da_2, db_0, db_1, db_2$ are the unknowns in the least square equations or the corrections for each transformation parameter; x_1, y_1 is the location of the pixel on the reference image; x_2^0, y_2^0 is the location which is determined by the geometric transformation parameters on the matched image; $\dot{g}_{2x}(x_2^0, y_2^0), \dot{g}_{2y}(x_2^0, y_2^0)$ are the derivatives along x and y directions at location (x_2^0, y_2^0) on the matched image, respectively. Once the initial location on the matched image or the initial transformation parameters are determined, equation 6.8 can be used for all the pixels in the window to solve for the unknowns iteratively. Since there are eight unknowns in equation 6.7, at least 8 equations are needed. So the window size should be larger than 3 by 3.

$$
\begin{aligned}
&g_2(x_2^0, y_2^0)dh_0 + h_0^0 \dot{g}_{2x}(x_2^0, y_2^0)da_0 + h_0^0 x_1 \dot{g}_{2x}(x_2^0, y_2^0)da_1 + h_0^0 y_1 \dot{g}_{2x}(x_2^0, y_2^0)da_2 \\
&+h_0^0 \dot{g}_{2y}(x_2^0, y_2^0)db_0 + h_0^0 x_1 \dot{g}_{2y}(x_2^0, y_2^0)db_1 + h_0^0 y_1 \dot{g}_{2y}(x_2^0, y_2^0)db_2 + dh_1 \\
&= g_1(x_1, y_1) - (h_0^0 g_2(x_2^0, y_2^0) + h_1^0) \\
&h_0 = h_0^0 + dh_0 \\
&h_1 = h_1^0 + dh_1 \\
&a_0 = a_0^0 + da_0 \\
&a_1 = a_1^0 + da_1 \\
&a_2 = d_2^0 + da_2 \\
&b_0 = b_0^0 + db_0 \\
&b_1 = b_1^0 + db_1 \\
&b_2 = b_2^0 + db_2 \\
&x_2^0 = a_0^0 + a_1^0 x_1 + a_2^0 y_1 \\
&y_2^0 = b_0^0 + b_1^0 x_1 + b_2^0 y_1
\end{aligned}
\qquad 6.8
$$

If the relative orientation of the stereo pair is good enough, and the stereo pair is in the epipolar coordinate system, the transformation parameters a_2, b_1 in equation 6.6 can be considered very small values and are ignored. Then the equation 6.8 becomes:

$$
\begin{aligned}
&g_2(x_2^0, y_2^0)dh_0 + h_0^0 \dot{g}_{2x}(x_2^0, y_2^0)da_0 + h_0^0 x_1 \dot{g}_{2x}(x_2^0, y_2^0)da_1 + h_0^0 \dot{g}_{2y}(x_2^0, y_2^0)db_0 \\
&+h_0^0 y_1 \dot{g}_{2y}(x_2^0, y_2^0)db_2 + dh_1 = g_1(x_1, y_1) - (h_0^0 g_2(x_2^0, y_2^0) + h_1^0) \\
&h_0 = h_0^0 + dh_0 \\
&h_1 = h_1^0 + dh_1 \\
&a_0 = a_0^0 + da_0 \\
&a_1 = a_1^0 + da_1 \\
&b_0 = b_0^0 + db_0 \\
&b_2 = b_2^0 + db_2 \\
&x_2^0 = a_0^0 + a_1^0 x_1 \\
&y_2^0 = b_0^0 + b_2^0 y_1
\end{aligned}
\qquad 6.9
$$

There are only six unknowns in the above equation. Therefore the time spent in computing will be less than time in computing the eight parameters equation 6.8.

6.3.2 Determination of the Initial Transformation Parameters in LSIM

Before the iterative process in solving equations 6.8 and 6.9, the initial values of these unknown transformation parameters must be estimated. These initial values can be determined by other image matching algorithms like cross correlation or least difference, or they can be obtained from the average ground height, and some assumption about the radiometric parameters can be made. First of all, before the first iteration, all the photographs in the stereo pair can be considered at the same radiometric scale and no radiometric shift among them, so the initial value of h_0, h_1 are:

$$\begin{aligned} h_0^0 &= 1 \\ h_1^0 &= 0 \end{aligned} \qquad 6.10$$

For the initial parameters from the average ground height, assume the rectification of these photographs are nearly perfect, all the photographs have the same scale and not much y parallax left, then the initial values of a_1, a_2, b_1, b_2 are:

$$\begin{aligned} a_1^0 &= 1 \\ a_2^0 &= 0 \\ b_1^0 &= 0 \\ b_2^0 &= 1 \end{aligned} \qquad 6.11$$

a, b are actually the shift from the location on the reference image to the estimated location on the image to be matched. Equation 6.12 gives their initial values as:

$$\begin{aligned} a_0^0 &= t \\ b_0^0 &= \frac{y_{01} - y_{02}}{C_y} \end{aligned} \qquad 6.12$$

Where t is determined by equation 5.4; y_{01}, y_{02} are the initial epipolar y photo coordinates for the reference and matched images, respectively; C_y is the y pixel size. Therefore, all the unknown parameters can be initialized by using equations 6.10, 6.11, 6.12.

If the initial matching position for LSIM is given by other image matching algorithms, then the initial pixel location and radiometric relationships are known. For example, assuming the initial pixel locations on the reference and matched images are x_1, y_1, x_2, y_2, the equation 6.11 is still valid and h_1 is 0, then

$$a_0^0 = x_2 - x_1$$
$$b_0^0 = y_2 - y_1 \qquad 6.13$$
$$h_0^0 = \frac{\bar{g}_2}{\bar{g}_1}$$

can be used to compute the initial values for a_0, b_0, h_0. Where $\bar{g}_1, \bar{g}_2$ are the mean pixel gray values in the window for both images.

6.3.3 Pixel Interpolating and Resampling

Since the pixel location x_2, y_2 calculated with equation 6.6 are no longer integers in the iterative process, the pixel gray values on these non-integer locations in the window have to be derived from the surrounding pixels by resampling and interpolation. If the interpolated pixels are put in a matrix called *win*(I, J) for $I = 0$ to $m+1$(number of columns); $J = 0$ to $n+1$ (number of rows) as the size of the window, where m, n are the number of columns and rows of the window used in the image matching, and bilinear interpolation which has been explained in equation 2.13 is used, then from interpolation each pixel in the matrix *win*(I, J) will have pixel gray value as:

$$i = \text{int}(x_2 - 0.5)$$
$$j = \text{int}(y_2 - 0.5)$$
$$I' = i - \frac{m}{2}$$
$$J' = j - \frac{n}{2} \qquad 6.14$$
$$win(I,J) = (1-y_2+j)\big((1-x_2+i)g_2(I+I',J+J')+(x_2-i)g_2(I+I'+1,J+J')\big)$$
$$+(y_2-j)\big((x_2-i)g_2(I+I'+1,J+J'+1)+(1-x_2+i)g_2(I+I',J+J'+1)\big)$$

Here i, j are the largest integers which are less than x_2, y_2, respectively. After interpolating with equation 6.14, all the pixels in the matrix *win*(I, J) can be used to generate the least square equations for the next iteration, therefore in the following discussion, the matrix *win*(I, J) is used instead of g_2. Notice that the matching window is two pixels less than the size of *win*(I, J), which ranges from I = 1 to m on column and J = 1 to n on row. Why the interpolated window is two pixels larger than the image matching window will be explained in the next section.

6.3.4 Determination of the Derivatives in x and y Directions

In equations 6.8 and 6.9, the derivatives in x and y of pixel location (x_2, y_2) on the matched image $\dot{g}_{2x}(x_2, y_2), \dot{g}_{2y}(x_2, y_2)$ are needed in order to solve the least square equations. As shown in Figure 6.1, point 0, 1, 2 are three adjacent pixels with the values of y_1, y_2, y_3. If a parabolic line $y = ax^2 + bx + c$ is used to connect the three pixels, the constants a, b, c in the equation are:

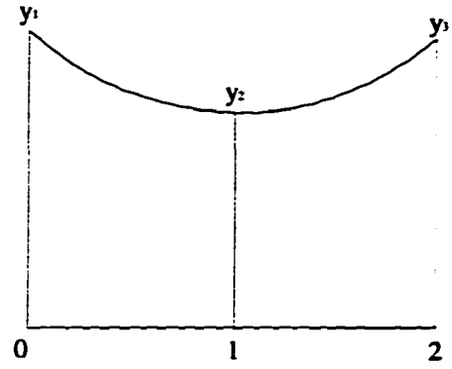


Figure 6.1 Determination of Tangent at Point 1

$$c = y_1$$

$$a = \frac{y_3 - 2y_2 + y_1}{2}$$

$$b = y_2 - y_1 - a$$

Then the slope at the point 1 is:

$$\dot{y}_1 = 2ax + b = 2a + b$$

Substituting a, b with y_1, y_2, y_3 the above equation becomes:

$$y_1 = \frac{y_3 - y_1}{2} \qquad 6.15$$

So the slope at a pixel in a given direction is the half difference of pixel gray values of the two pixels surrounding it in that direction. For those pixels on the edge of the matching window, one more pixel outside the matching window boundary is needed to compute the

slopes of the edge pixels. This is why the interpolated window is two pixels larger than the matching window.

6.4 Searching Strategies: Whole Range and Best-track

Two types of searching strategies, whole range and best-track, are used in this study.

6.4.1 Whole Range Searching

In whole range searching, all the pixels in the searching range in both rows and columns are used in the matching. As shown in Figure 6.2, once the initialization is done as described

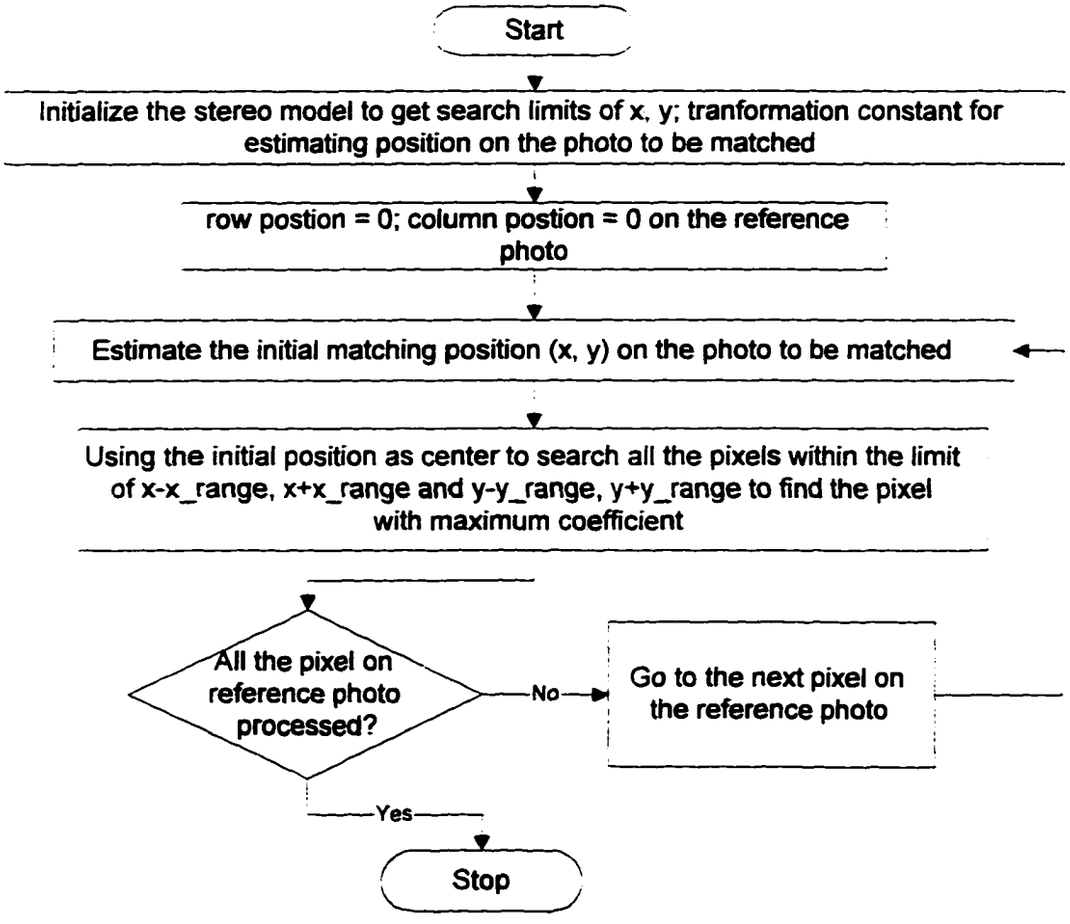


Figure 6.2 Flowchart of Whole Range Searching Algorithm

in Chapter 5, for each pixel on the reference photograph, all the pixels on the photo to be matched within the defined search area will be used in the cross correlation process to obtain the location with the maximum cross correlation coefficient.

In Figure 6.2, x_range, y_range are used to define the searching range on each row and column. The initialization process was described in detail in Chapter 5. As mentioned before, the basic idea for the whole range searching is that all the pixels within the defined range will be used in the matching process.

6.4.2 Best-track Searching

In the best-track searching, only the pixels on certain rows and columns are used in the matching, as shown in Figure 6.3.

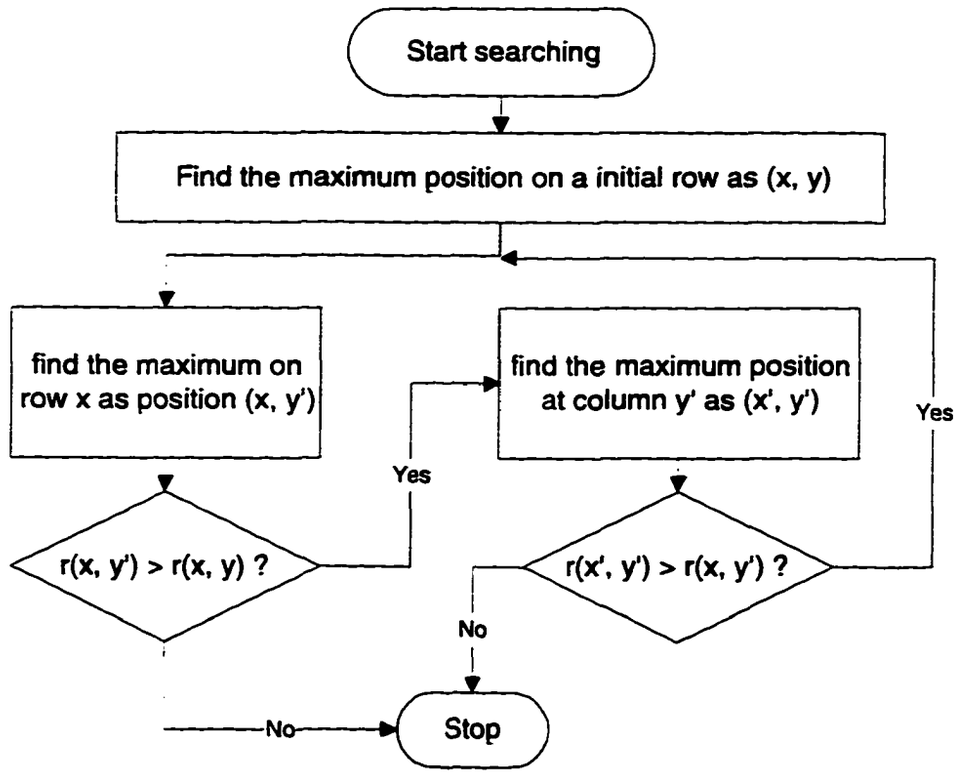


Figure 6.3 Flowchart of Best-track Searching Algorithm

The idea of the best-track searching is to follow the positions with maximum coefficient on the rows and columns until the largest is found. So the basic steps in the searching is to find the position on a row or column with the largest coefficient, then continue to search the column or row from that position, and repeat the same process within the search limit until the largest is located.

It is obvious that this searching algorithm will be faster than the whole range searching algorithm. They will be compared in the next section.

6.4.3 Comparison between These Two Types of Searching Strategies

The comparison is done in terms of speed and accuracy. As shown in Figure 6.4 a, it can be seen that the best-track searching algorithm is much faster, on the average the best-track is almost three times as fast as the whole range searching.

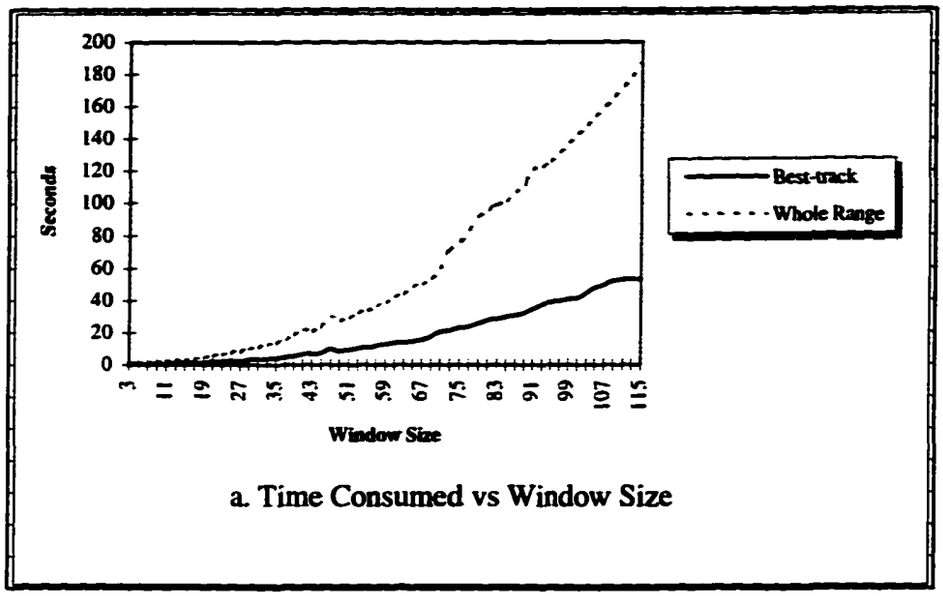
Figure 6.4 b shows the precision comparison of these two methods. When the window size is less than 35 by 35, the whole range searching is obviously better than the best-track searching, but there is almost no difference in the precision when the window size is larger than 35 by 35.

Therefore, the conclusion here is the best-track searching algorithm is better than the whole range if the matching window size is big enough. In this research case, the minimum window size is about 35 by 35.

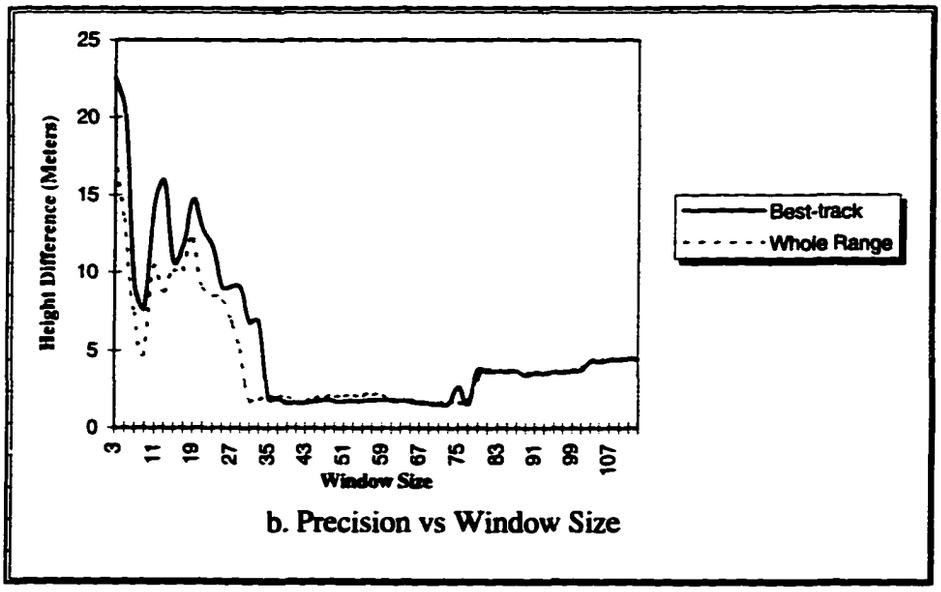
6.4.4 Comparison between x and y First Best-track Searching

The best-track searching algorithm shown in Figure 6.3 starts with x axis or rows, which is called x first best-track searching. In the best-track image matching algorithm, either x or y first searching can be used.

The comparisons between them on matching speed and precision are shown in Figure 6.5 a, b. These two figures show that they have no much difference on the time consumed, but the precision difference between them is significant when window size is less than about 47 by 47, the x first best-track searching has the better matching precision. So all the best-track searching used in rest of the study refers to the x first best-track searching.



a. Time Consumed vs Window Size



b. Precision vs Window Size

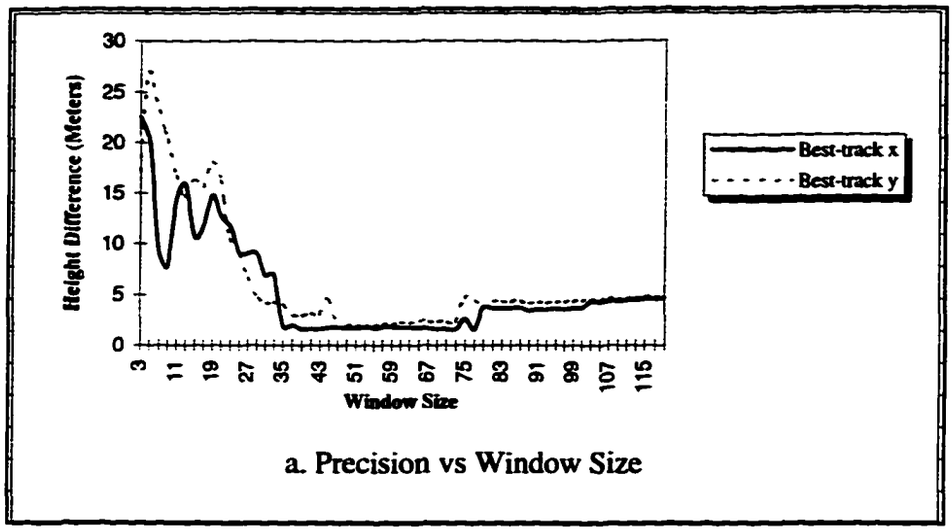
Figure 6.4 Precision and Speed Comparison between Best-track and Whole Range

6.5 Comparison between CCIM and LDIM

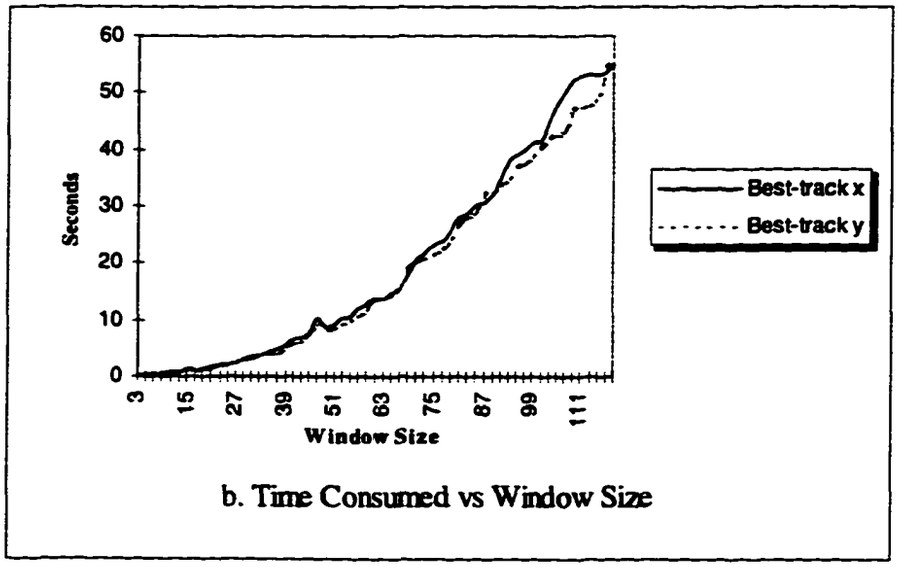
Generally speaking the CCIM has higher precision than LDIM but it is slower. In this section the empirical matching characteristics of these two image matching methods are discussed and compared using the search algorithms.

6.5.1 Correlation Coefficient Surface from CCIM

Figure 6.6 shows the correlation coefficient surface from CCIM. The vertical axis in the Figure is the coefficient values; the x axis is the number of columns from the initially estimated position on the image to be matched; the y axis is the number of rows from the initially estimated position on the image to be matched. The width in the y direction is given by the search range in y direction. The window size used here is 37 by 37 which gives the good matching results in Figure 6.5 and in others shown later in this chapter.



a. Precision vs Window Size



b. Time Consumed vs Window Size

Figure 6.5 Comparison between x and y First in Best-track

From Figure 6.6 it is obvious that the control points and the ground points have the best correlation surface shape whose mountain peak gives a good location of the maximum and the coefficient surface is more smooth. There is no big slope change in the correlation coefficient surface from the control points and the ground points. The surface shape from tree and grass points and the roof points of building are not so good because the correlation coefficients change a lot in some areas, especially for the roof points. The grass and tree points are better than the roof points since its peak location is more obvious and their surface is better than the roof points; they also have more discontinuous surrounding features even though they may have better contrast. Therefore the points with distinct and continuous surrounding features will have the best matching results from CCIM.

Since there are four different types of ground features, for a point from a certain type of ground feature, the correlation coefficient surface from CCIM at different window size are shown in Figures 6.7, 6.8, 6.9 and 6.10.

The first point is the control point No. 1720. This is a control point in a center of a sidewalk surrounded by grass. It is a point with good surrounding contrast and continuity. As shown in Figure 6.7, the first matching window size is 3 by 3. The interesting thing is that the correlation coefficient surface for the window of 3 by 3 looks just like the intersection on the photographs. This also means the information available in the image matching is not sufficient. The area outside the edge of the sidewalk has smaller correlation coefficients which makes the sidewalks appear on the coefficient surface. The sidewalk appears on the coefficient surface until the window size is about 27 by 27. On the other hand, the maximum of the coefficient surface rises and becomes obvious with respect to the surroundings as the window size increases. Therefore, for a point with good contrast and continuity to its surrounding, the CCIM can give good matching results when the matching window size is larger than a certain number, such as 11 by 11, in this study case.

The second point is also a point on the ground, point No. t15. It is a corner of sidewalk with tress and buildings nearby. In a small area it has good contrast, but the trees and buildings bring a lot of surrounding discontinuity.

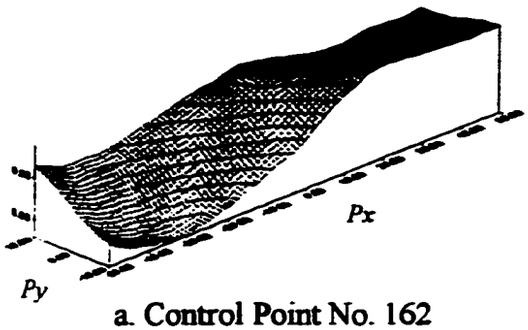
Like the first point, the coefficient surface from the 3 by 3 window size also shows the shape of the ground feature. It is hard to tell where the peak is even at a window size of 11 by 11. The location of the peak appears clearly in the window size of 19 by 19. The second point has the best coefficient surface shape when the window size is between 27 by 27 and 51 by 51.

After the window size increases beyond 75 by 75, the coefficient surface worsens losing the peak because the discontinuous surroundings begin to be involved and this affects the matching results. When the window size is between 19 by 19 and 51 by 51, the matching is good because of the good contrast in the small area. but the matching worsens when the window size is larger than 51 by 51 because more and more discontinuity surrounding is involved. From the above discussion the conclusion is that for points with some good contrast and some discontinuity surroundings, the better the good contrast portion is in the searching window, the better the matching result.

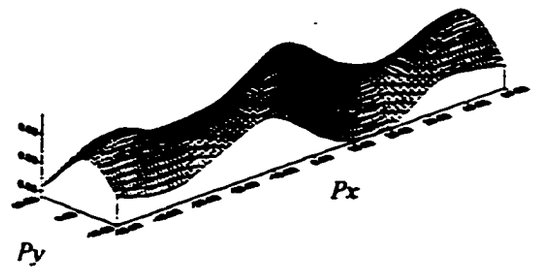
Figure 6.9 shows the correlation coefficient surface of the roof point No. t33 at different window size. Actually this point is the top of the water tower. It is the highest point in the ISU central campus area. So it has the most discontinuity surrounding, and it should have the worst matching result. According to its height, the matching location should be located on the left-hand side. Windows of 3 by 3 and 11 by 11 do not give the right matching results. They give the right location of the maximum when the window size is between 35 by 35 and 75 by 75, even though the surface shape around the maximum is not good. It seems that the discontinuity here causes too many variations.

Figure 6.10 shows the correlation coefficient surface for the point No. t24 which is a tree and grass point. This point is on top of a small bush within a lawn, with tree and grass and a little low contrast. From Figure 6.10 we see that it does not have a good correlation coefficient surface until the window size is 19 by 19. Since there is not much discontinuity in this area, the larger the window size, the better the matching .

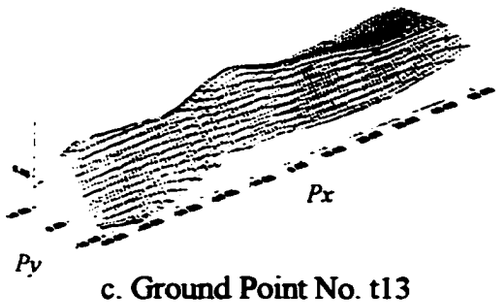
From the above discussions on these correlation coefficient surface we see that the trees and grass have low contrast and the buildings bring discontinuity to the matching.



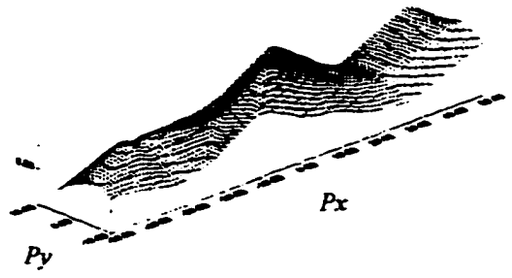
a. Control Point No. 162



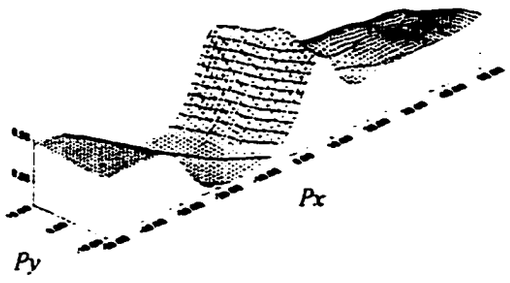
b. Control Point No. 1720



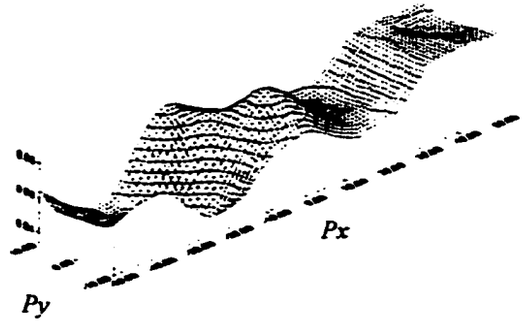
c. Ground Point No. t13



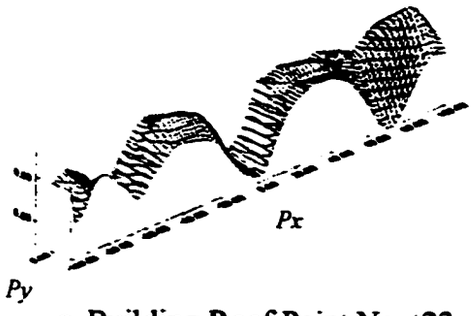
d. Ground Point No. t15



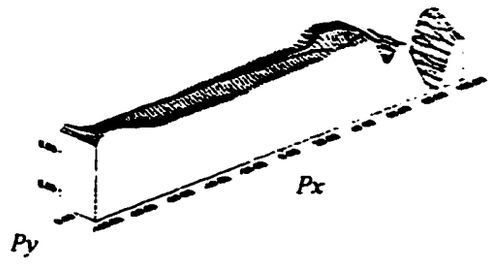
e. Tree and Grass Point No. t23



f. Grass and Tree Point No. t24

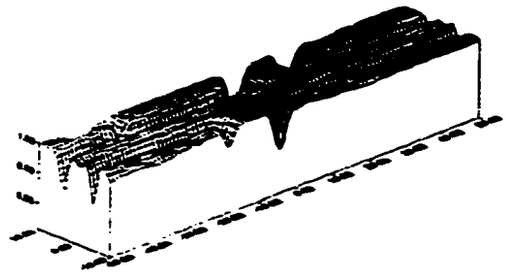


g. Building Roof Point No. t33

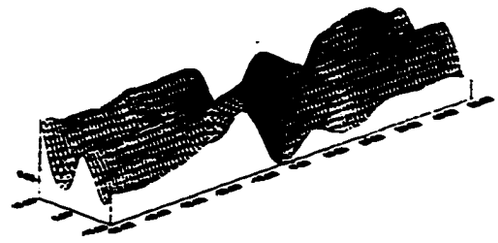


h. Building Roof Point No. t34

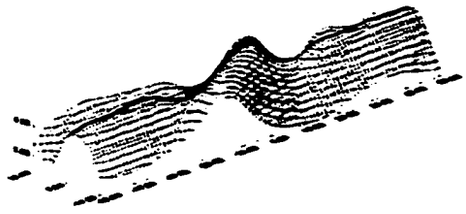
Figure 6.6 Correlation Coefficient Surface from CCIM
Window Size: 27 by 27



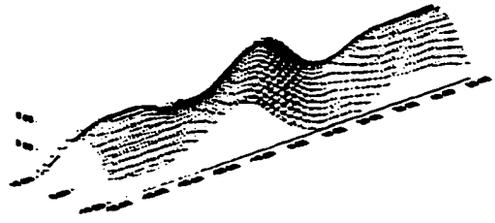
a. 3 by 3



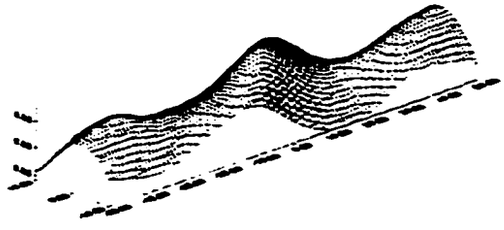
b. 11 by 11



c. 19 by 19



d. 27 by 27



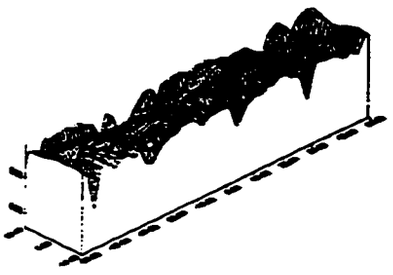
e. 35 by 35

f. 51 by 51

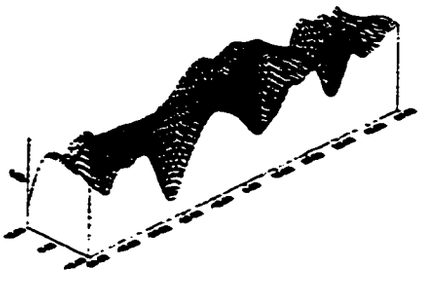
g. 75 by 75

h. 83 by 83

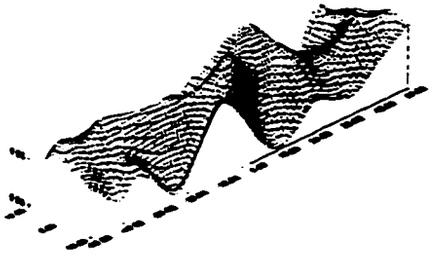
Figure 6.7 Correlation Surface for Control Point No. 1720
at Different Window Size



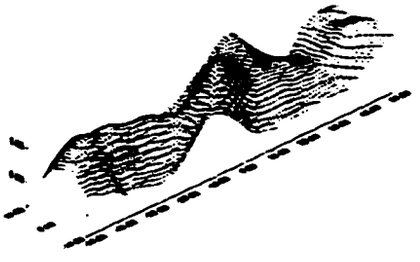
a. 3 by 3



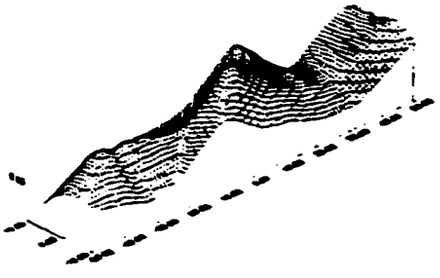
b. 11 by 11



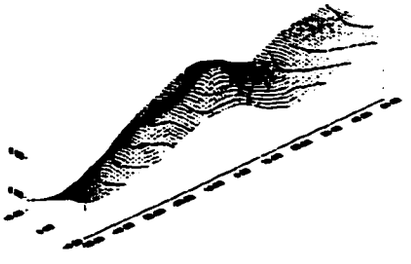
c. 19 by 19



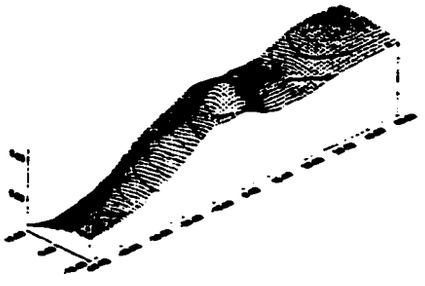
d. 27 by 27



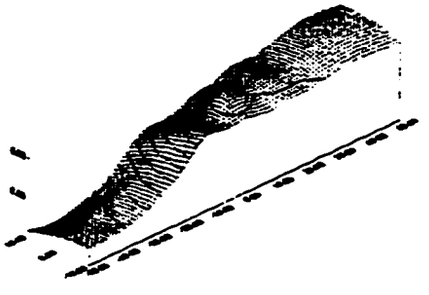
e. 35 by 35



f. 51 by 51

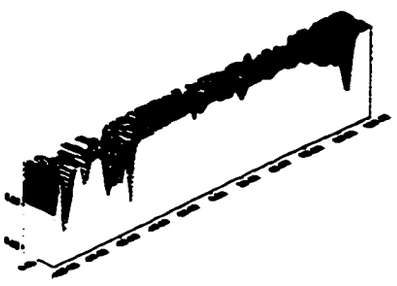


g. 75 by 75



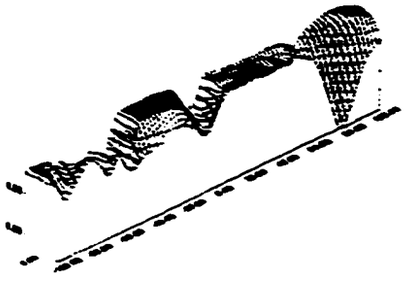
h. 83 by 83

Figure 6.8 Correlation Coefficient Surface at Different Window Size for Ground Point No. t15

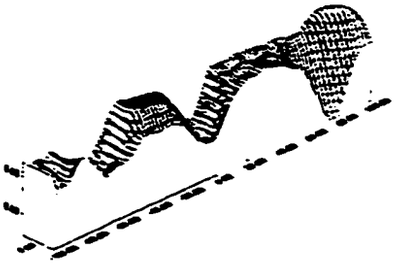


a. 3 by 3

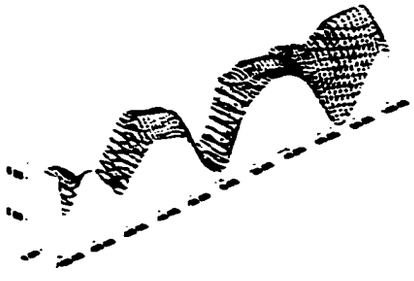
b. 11 by 11



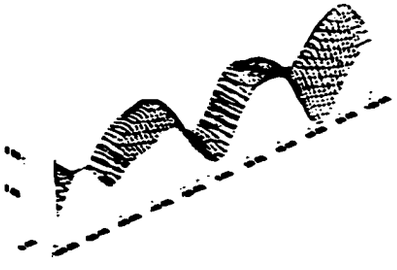
c. 19 by 19



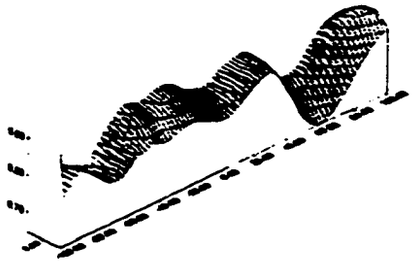
d. 27 by 27



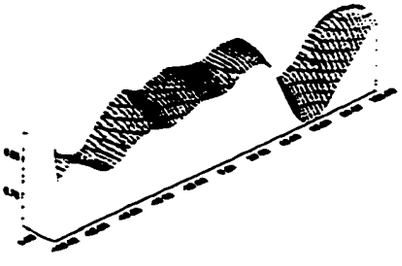
e. 35 by 35



f. 51 by 51

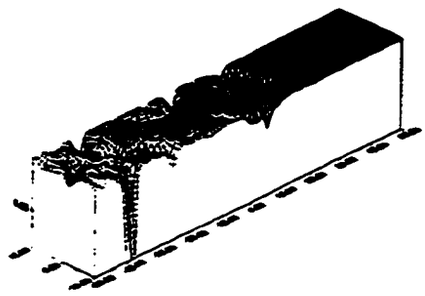


g. 75 by 75

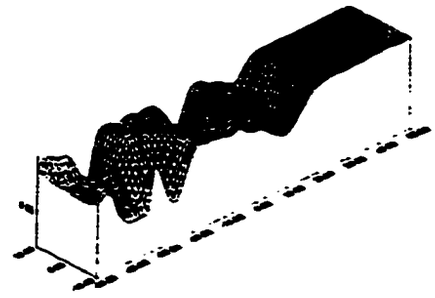


h. 83 by 83

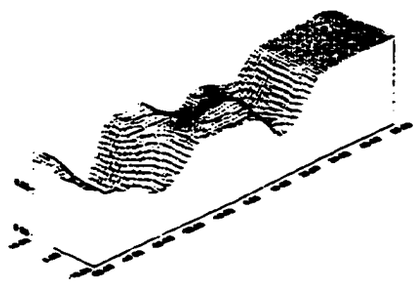
Figure 6.9 Correlation Coefficient Surface at Different Window Size for Building Roof Point No. t33



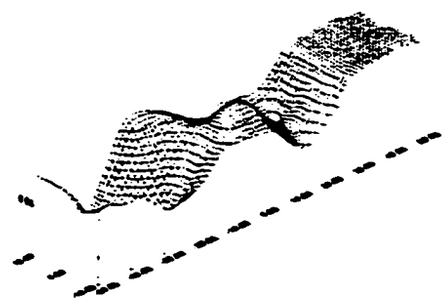
a. 3 by 3



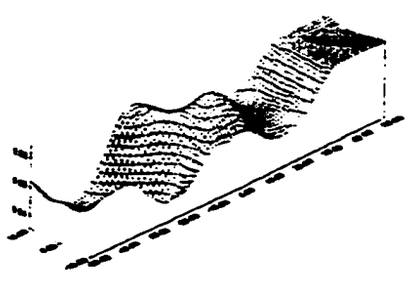
b. 11 by 11



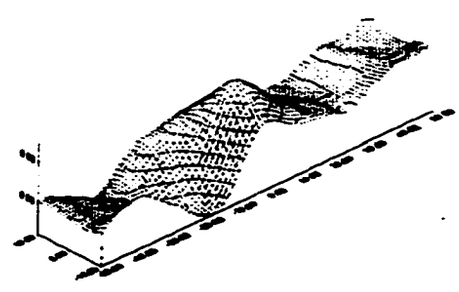
c. 19 by 19



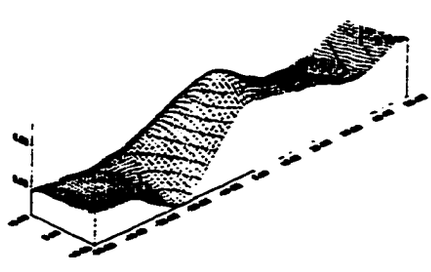
d. 27 by 27



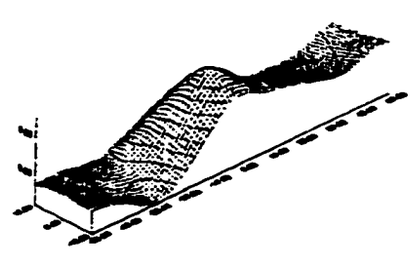
e. 35 by 35



f. 51 by 51



g. 75 by 75



h. 83 by 83

Figure 6.10 Correlation Coefficient Surface at Different Window Size for Tree and Grass Point No. t24

The larger window size is good for low contrast area, but it is bad for discontinuous ground features.

6.5.2 Difference Surface from LDIM

Figure 6.11 shows the difference surface from LDIM. As in CCIM, the window size used in the following figures is also 37 by 37. This is good for the comparison.

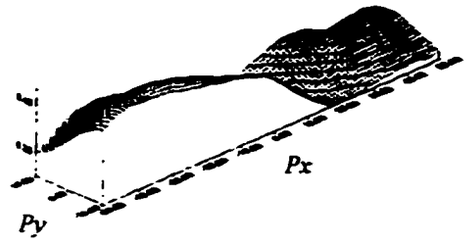
The difference surface from eight points of four different types of ground features are shown in the Figure 6.11. Note that unlike in CCIM, the minimum value on the surface is used in LDIM to locate the matching position.

For a given point, different window sizes give various difference surface, and the results are compared. Figures 6.12, 6.13, 6.14, 6.15 give the difference surface from LDIM vs. window size for each type of ground feature. The same points are used here as those in the previous section for comparison.

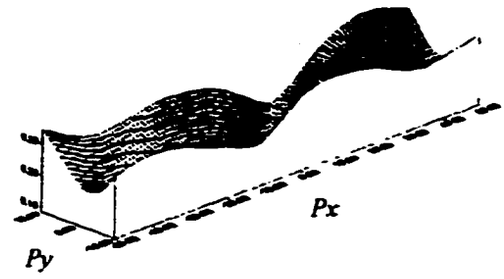
Figure 6.12 shows the difference surface from control point No. 1720. The difference surface gives good results at the window size of 3 by 3 even though the surface is rough. It is obvious that the larger the window size, the better the matching because the surface gives better location of the minimum as the window size increases. This is because the point is on an open area with good contrast and low discontinuity. The LDIM gives the same matching results as the CCIM does for this point.

Figure 6.13 gives the difference surface for the point No. t15. Obviously the window of size 3 by 3 does not give good matching result. Although window of 11 by 11 has a much better result than window of 3 by 3, it is still not good enough. The matching gives good result when window size is larger than 19 by 19. The best results are when the window size is between 35 by 35 and 51 by 51, beyond that, the difference surface begins to worsen, and finally at window of 83 by 83 some portion is eliminated because of the bad results from LDIM. The conclusion is similar to that from CCIM.

Figure 6.14 gives the difference surface form the point No. t33. Since this point is the highest point in this area, has large height difference, and is also surrounded by a lot of trees and buildings, so it should have the largest discontinuity. This can be seen on the difference surface at the window sizes of 3 by 3 and 11 by 11. The matching position should be on the

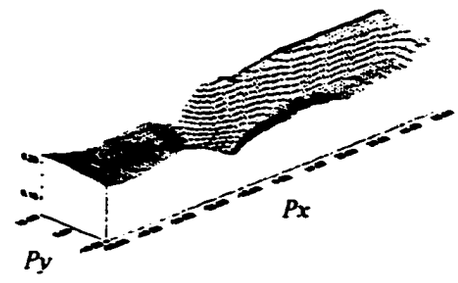


a. Control Point No. 162

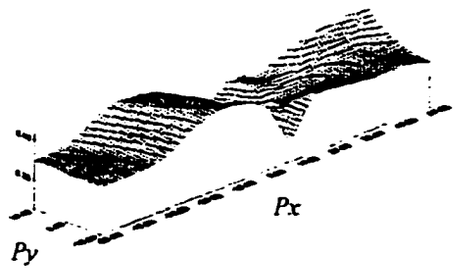


b. Control Point No. 1720

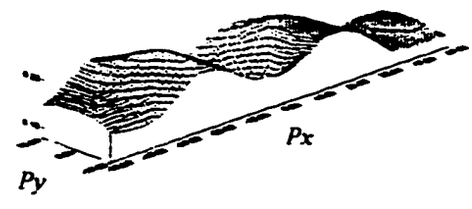
c. Ground Point No. t13



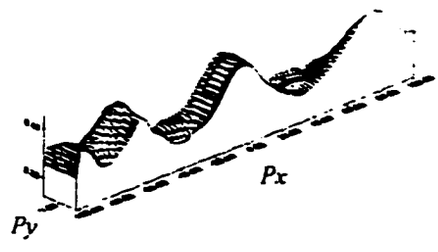
d. Ground Point No. t16



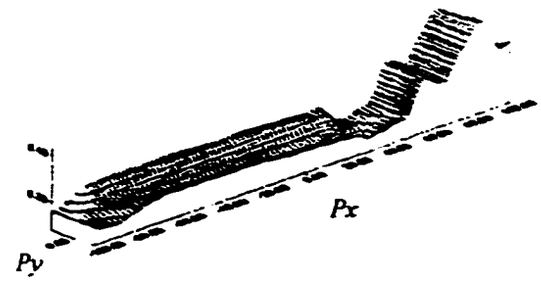
e. Tree and Grass Point No. t23



f. Tree and Grass Point No. t24

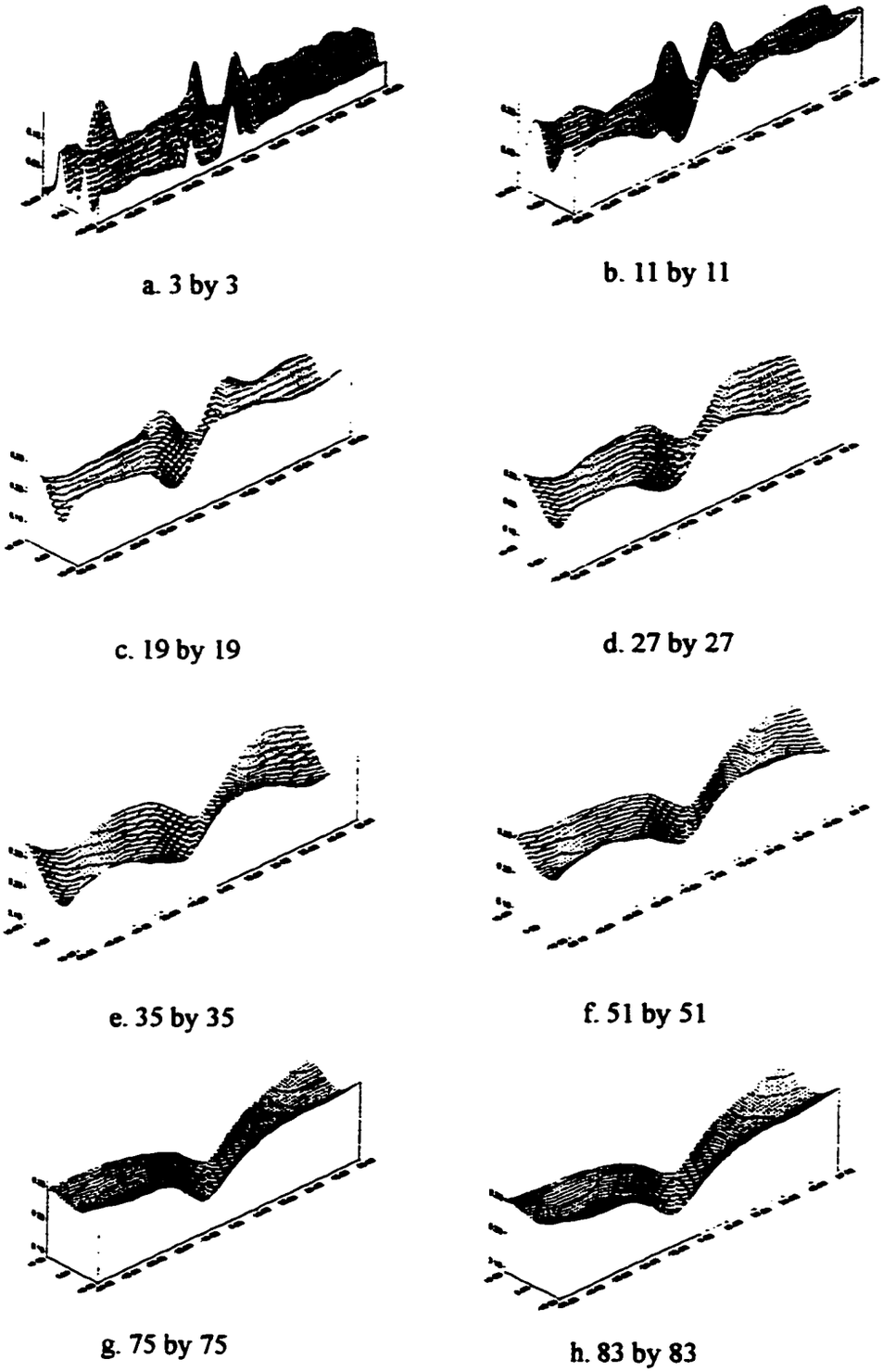


g. Building Roof Point No. t33



h. Building Roof Point No. t34

Figure 6.11 Difference Surface from LDIM
Window Size: 27 by 27



a. 3 by 3

b. 11 by 11

c. 19 by 19

d. 27 by 27

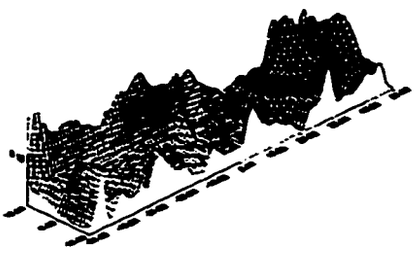
e. 35 by 35

f. 51 by 51

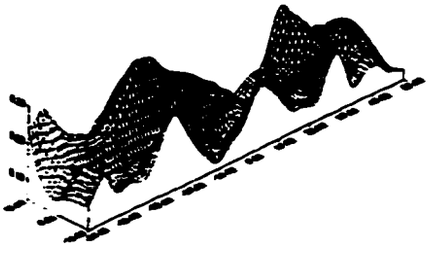
g. 75 by 75

h. 83 by 83

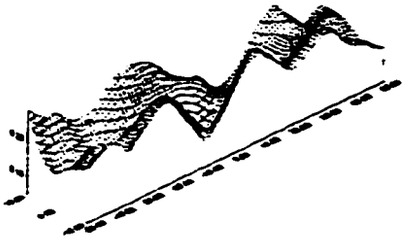
Figure 6.12 Difference Surface at Different Window Size
for Control Point No. 1720



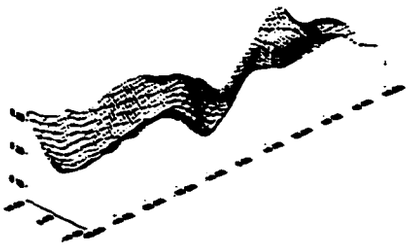
a. 3 by 3



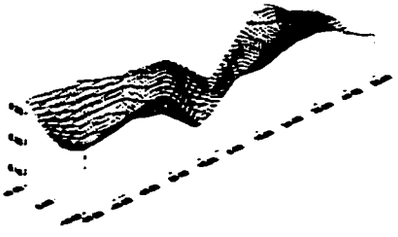
b. 11 by 11



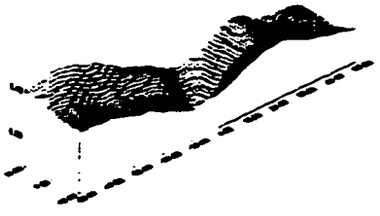
c. 19 by 19



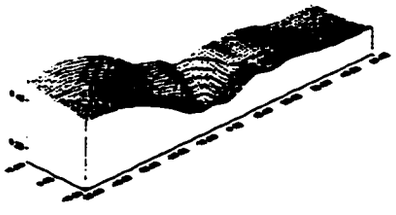
d. 27 by 27



e. 35 by 35



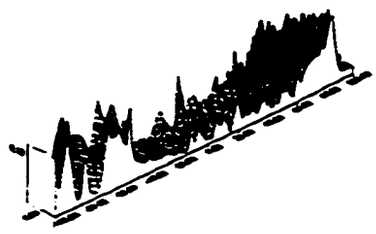
f. 51 by 51



g. 75 by 75

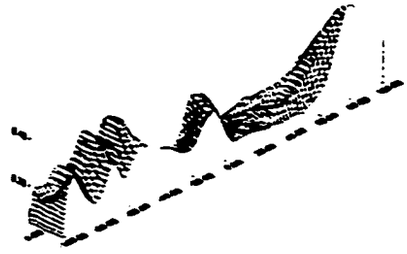
h. 83 by 83

Figure 6.13 Difference Surface at Different Window Size
for Ground Point No. t15

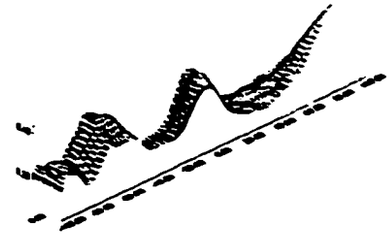


a. 3 by 3

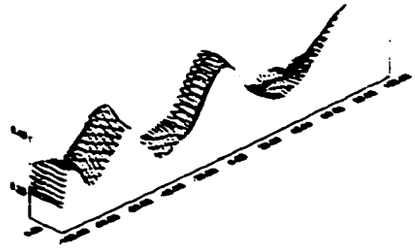
b. 11 by 11



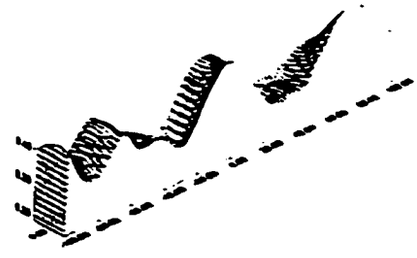
c. 19 by 19



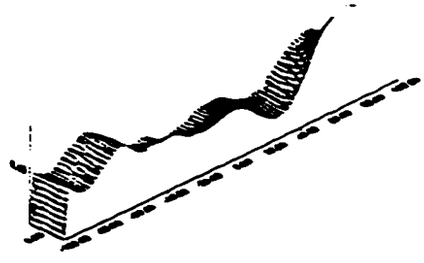
d. 27 by 27



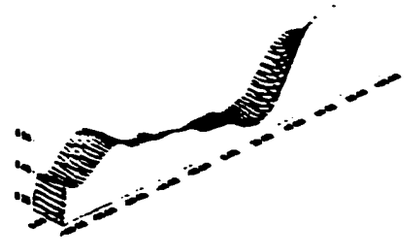
e. 35 by 35



f. 51 by 51

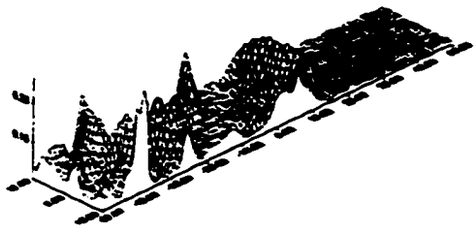


g. 75 by 75

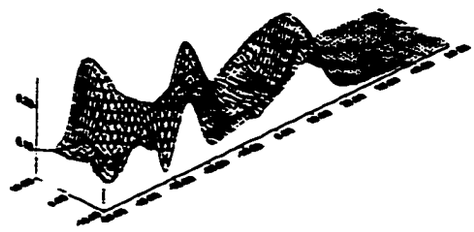


h. 83 by 83

Figure 6.14 Difference Surface at Different Window Size
for Building Roof Point No. t33



a. 3 by 3

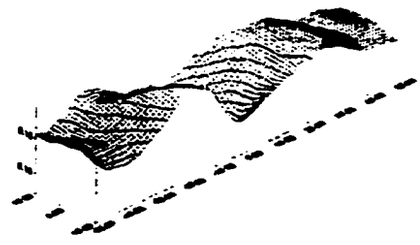


b. 11 by 11

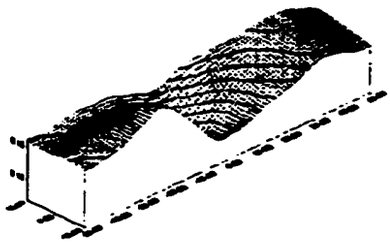
c. 19 by 19

d. 27 by 27

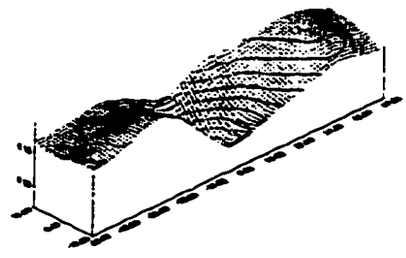
e. 35 by 35



f. 51 by 51



g. 75 by 75



h. 83 by 83

Figure 6.15 Difference Surface at Different Window Size
for Tree and Grass Point No. t24

left-hand side. The difference surface is too rough to indicate where the minimum value is. It is interesting that the window of 19 by 19 gives the right matching location which is located on the left-hand side, but the windows of 27 by 27 and 35 by 35 do not give the right matching locations. However when the window size is larger than 51 by 51 we again get good results. This is because the discontinuity causes inconsistent results. Therefore LDIM is less sensitive to the discontinuity than the CCIM.

The difference surface for point No. t24 is shown in Figure 6.15. The matching result is not good until the window size is larger than 35 by 35. The matching result is better when window size is between 51 by 51 and 75 by 75. However, when window size is larger than 75 by 75, even though the matching result is good, the difference surface shape begins to worsen with the increase of the window size. It seems that the LDIM is more sensitive to the low contrast than the CCIM.

The conclusions from the above discussion are:

- Good contrast is needed for a good matching
- CCIM is more sensitive to discontinuity
- LDIM is more sensitive to low contrast

6.5.3 Comparison between CCIM and LDIM in Terms of Speed

Figure 6.16 shows the comparison between CCIM and LDIM in terms of speed for all 24 check points.

The difference between them is not much. Since much time is consumed in normalizing the results in LDIM, it seems that the LDIM is slower than CCIM.

6.5.4 Comparison between CCIM and LDIM in Terms of Accuracy

Figure 6.17 shows the accuracy difference between LDIM and CCIM for all 24 check points. There is much difference when the window size is less than 31 by 31, or between 71 by 71 and 81 by 81, and these two methods match well when the window size is between 31 by 31 and 71 by 71.

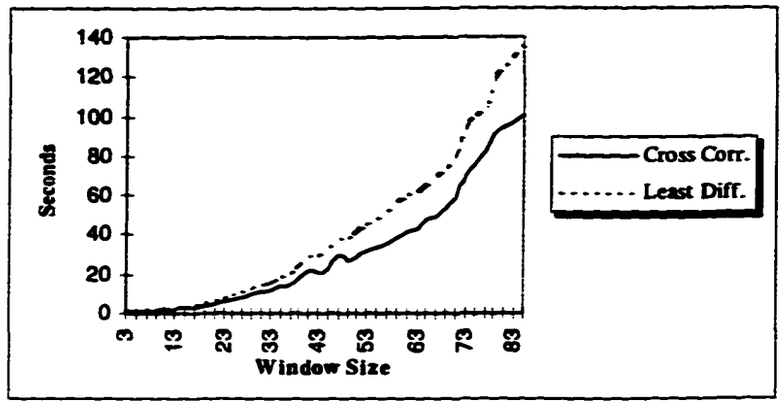


Figure 6.16 Time vs. Window Size for Cross Correlation and Least Difference

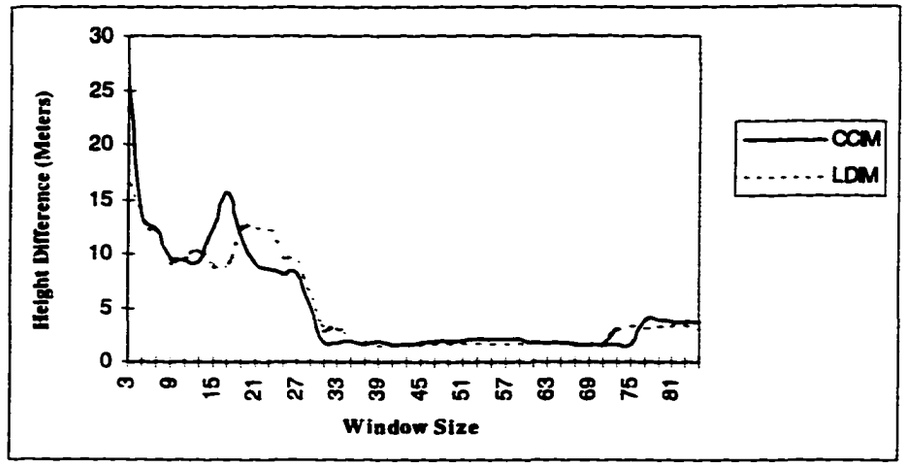


Figure 6.17 Accuracy vs. Window Size between CCIM and LDIM

6.5.5 Comparison between CCIM and LDIM in terms of Accuracy for Different Types of Ground Features

The comparison of accuracy for the four different types of points with CCIM is shown in Figure 6.18. According to the Figure 6.18 the roof points of buildings have the worst matching accuracy, and the ground points have the best results because they are the corners or intersections which can be correlated with no difficulty.

The comparison of different ground features with LDIM is shown in Figure 6.19. It has similar results as CCIM except that ground points from LDIM are better.

6.6 Image Matching Results from LSIM

The LSIM results mainly depend on the initial matching position. Here CCIM, LDIM and the estimate position from the average ground height are used as the initial matching position. As mentioned before, 8 or 6 parameters can be used in LSIM, and different number of iterations can be applied according to the real situation.

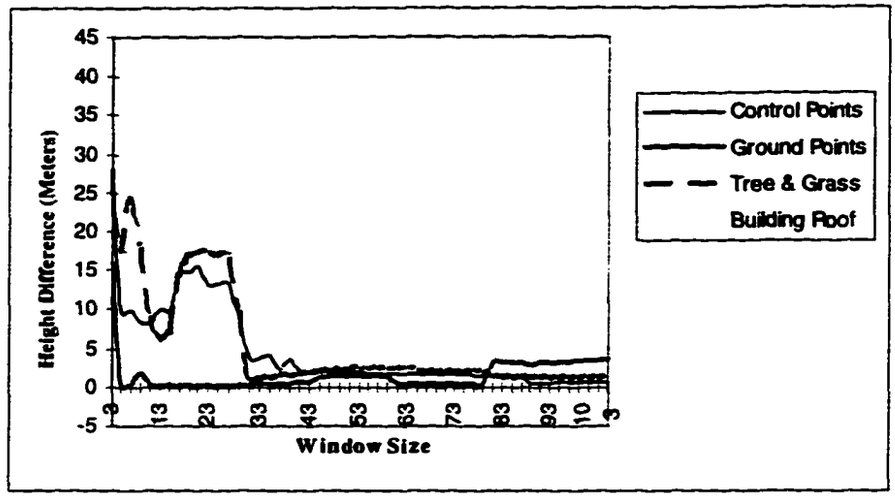


Figure 6.18 Accuracy vs. Window Size for Different Types of Ground Features with CCIM

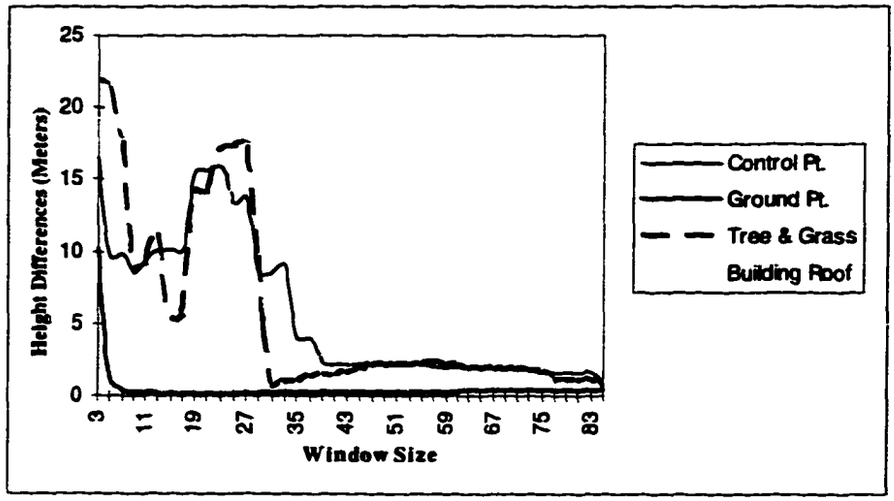


Figure 6.19 Accuracy vs. Window Size for Different Types of Ground Features with LDIM

6.6.1 LSIM with Initial Position from the Average Height

Figures 6.20 and 6.21 shows the comparison between 8 parameters and 6 parameters and 3 and 5 iterations. From Figure 6.20 it is obvious that the combination of 6 parameters and 3 iterations is the fastest, and the one with 8 parameters and 5 iterations is the slowest. From Figure 6.21 it is known that there is not much difference between 6 or 8 parameters and 3 or 5 iterations except that the 6 parameters have very a height difference around the window size of 71 by 71.

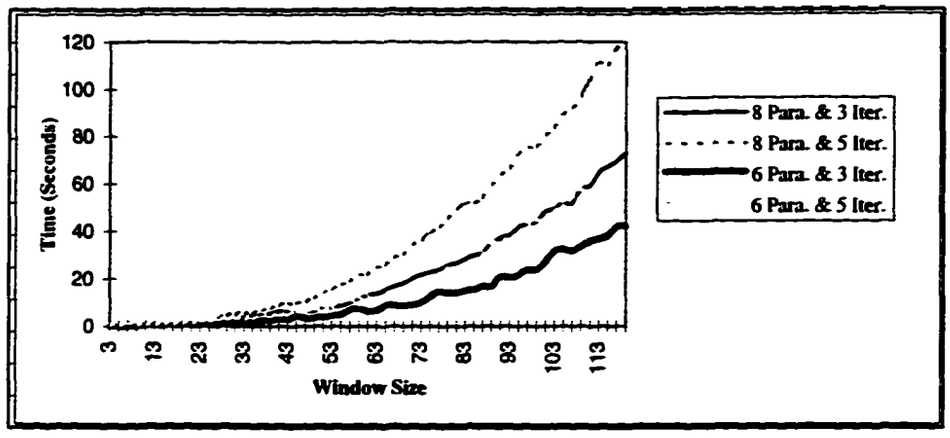


Figure 6.20 Comparison of Speed among 8 and 5 Parameters and 3 and 5 Iterations

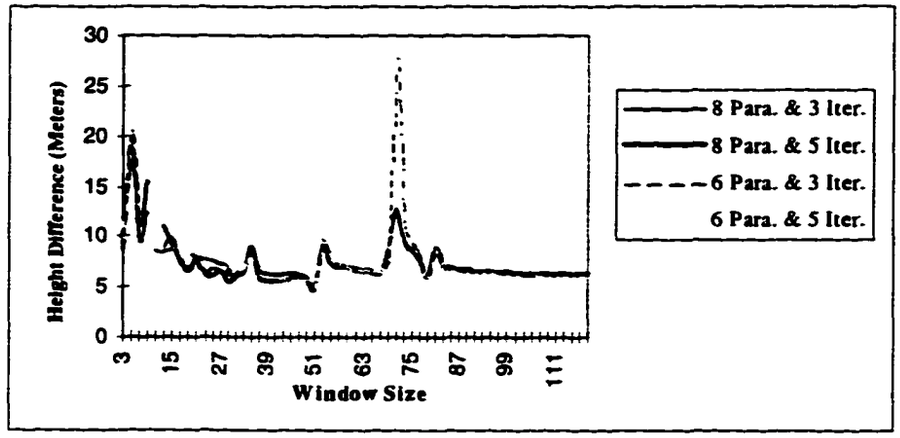


Figure 6.21 Comparison of Accuracy among 8 and 5 Parameters and 3 and 5 Iterations

So 8 parameter method is better than the 6, and 3 iteration is enough if the solution is convergent. The average accuracy here is about 6 meters which is worse than LDIM and CCIM. From the above results we see that the combination of 6 parameters and 3 iteration is good in most cases.

Since Figures 6.20 and 6.21 only give the mean results of the different ground features, in order to know more about each type of ground feature, we show the comparison between them in Figure 6.22. It is known that the roof points have the worst matching accuracy probably caused by the discontinuity of the surrounding. The height differences are big for the grass and tree points when the window size is less than 7 by 7. After the window size is larger than about 39 by 39, the matching results stabilize around 2 meters except for the roof points. Therefore the LSIM is not good for the image matching when there is a great deal of feature dissimilarities around the points on the photographs in the stereo pair.

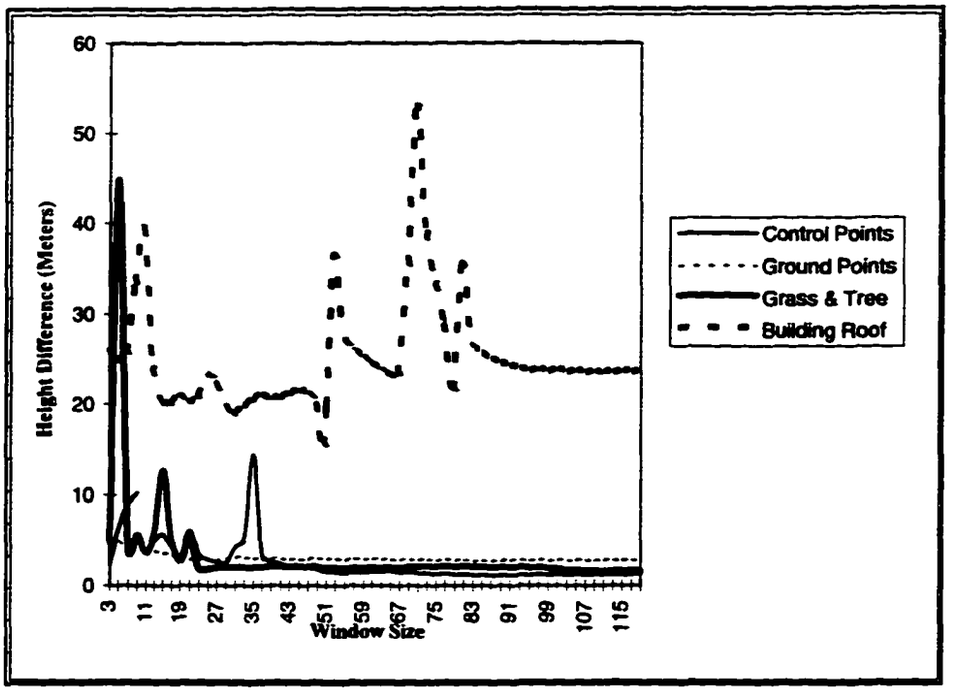


Figure 6.22 Comparison among Different Types of Ground Features (8 Para. and 3 Iter.)

6.6.2 LSIM with Initial Position from CCIM

Figure 6.23 shows the LSIM of 8 parameters and 3 iterations using the initial position from CCIM. It looks more like the results from CCIM.

6.6.3 LSIM with Initial Position from Average Ground Height

Figure 6.24 shows the LSIM results of 8 parameters and 3 iterations using the initial position from LDIM. Again it looks more like the results directly from LDIM.

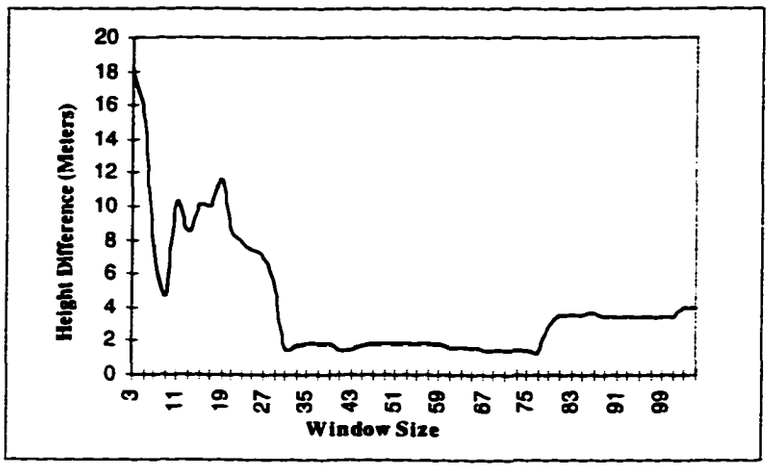


Figure 6.23 Accuracy of LSIM (8 Para. and 3 Iter.) with Initial Positions from CCIM

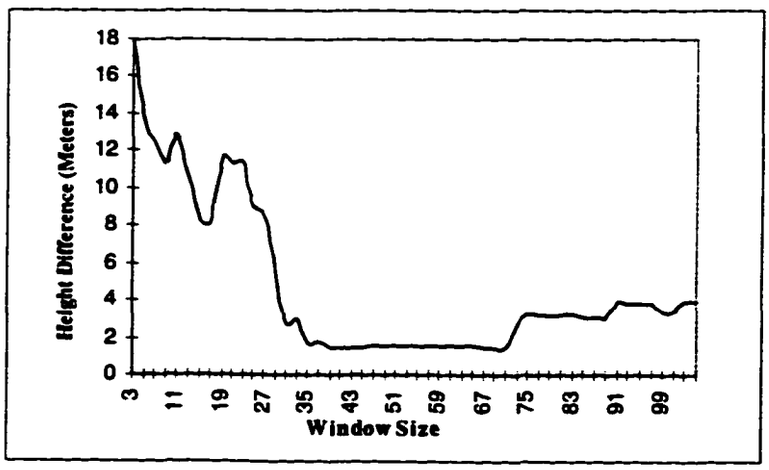


Figure 6.24 Accuracy of LSIM (8 Para. and 3 Iter.) with Initial Positions from LDIM

6.6.4 Comparison between the Results in Sections 6.6.1, 6.6.2 and 6.6.3

As shown in Figure 6.25, LSIM method with 8 parameters and 3 iterations has the fastest matching speed among these three image matching methods. LSIM is time consuming if the solution is not convergent or stable.

As shown in Figure 6.26, LSIM with the estimate from the average height is not as good as the CCIM and LDIM.

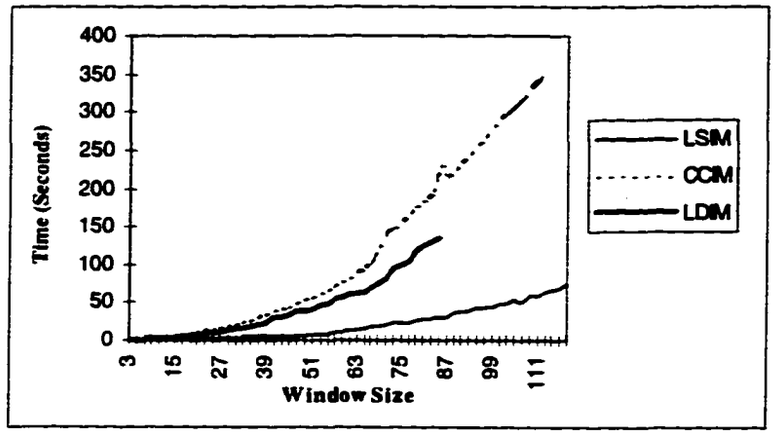


Figure 6.25 Speed Comparison among LSIM (8 Para. and 3 Iter.), CCIM and LDIM

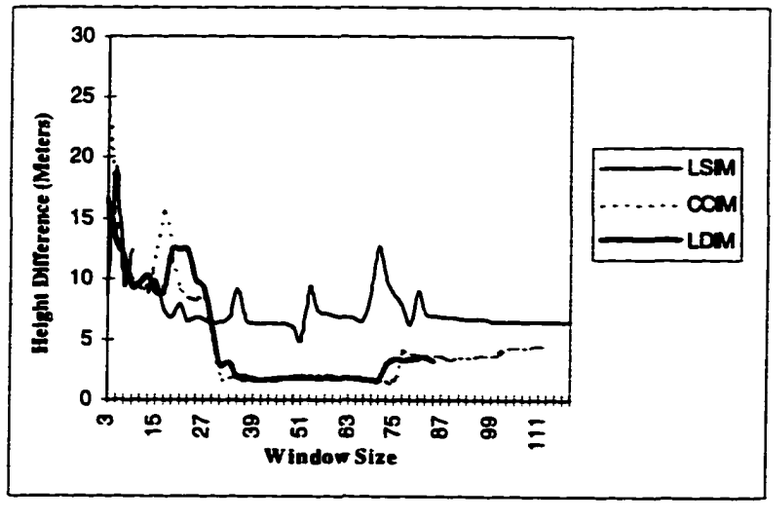


Figure 6.26 Accuracy Comparison among LSIM (8 Para. and 3 Iter.), CCIM and LDIM

Figure 6.27 shows the comparison between the CCIM and the LSIM using the initial position from CCIM. From this figure it is seen that there is not much change after applying LSIM to CCIM. The comparison between the LDIM and the LSIM using the initial position from LDIM shown in Figure 6.28, does not show much improvement after applying the LSIM to the LDIM. Compared to the LSIM results with the initial position from the average height, different initial positions let LSIM give out different results. Therefore LSIM is very sensitive to the initial location, and if the initial position is good, there is not much improvement from LSIM.

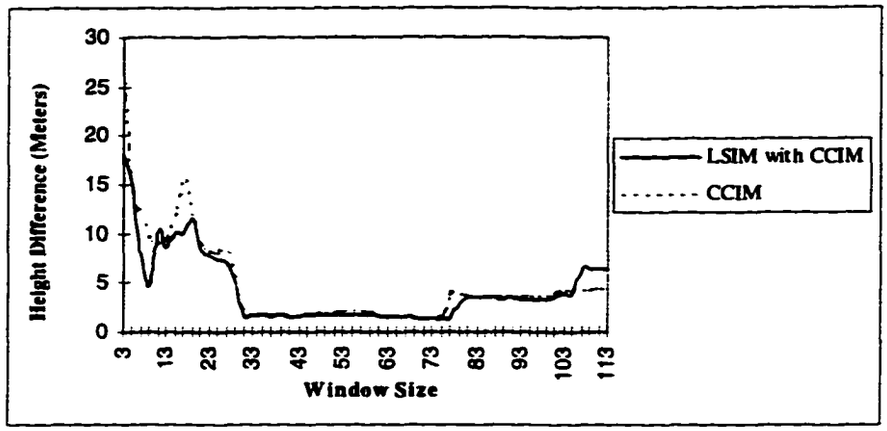


Figure 6.27 Comparison between LSIM (8 Para. and 3 Iter.) with CCIM and CCIM

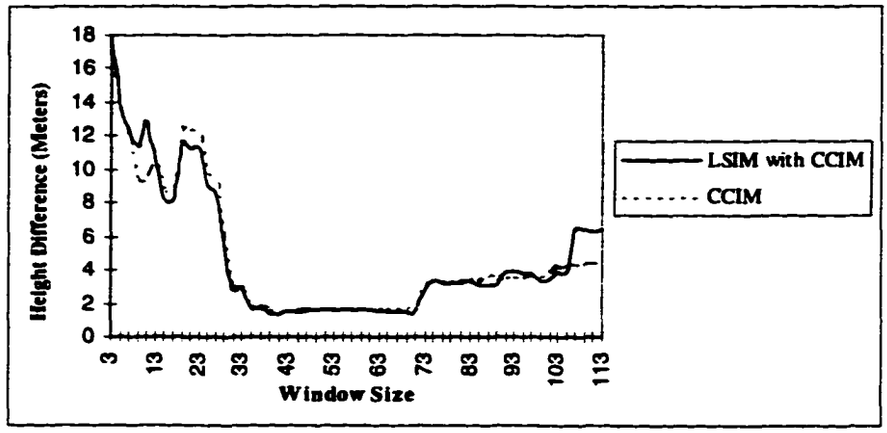


Figure 6.28 Comparison between LSIM (8 Para. and 3 Iter.) with LDIM and LDIM

6.7 Conclusions

From the above discussion on the three types of image matching algorithms, we can conclude:

- LSIM is the fastest of the three methods tested
- LSIM is sensitive to the initial position
- The results from CCIM or LDIM can be used as the initial position of LSIM, but there is not much difference from LSIM
- From the results shown in Figure 6.22, the building roof points have the lowest accuracy. Therefore LSIM is sensitive to the discontinuity of the feature on the photograph
- The roof points have the lowest accuracy in LDIM and CCIM, So discontinuity on the photographs can cause errors in the image matching.
- There is a range of window size which has the best matching results. In this case, the range of window size is from approximately 31 by 31 to 71 by 71.
- Good image matching needs (a). sufficient information for the matching (b). good contrast, (c). and little discontinuity

CHAPTER 7 MULTI-METHOD IMAGE MATCHING

The conclusion of the previous chapter is that different image matching algorithms differ in sensitivity to the same types of error on the photographs. Since the errors in different types of image matching methods even when they are from the same data source can differ, the comparison between them could show some of the errors. Therefore this method can be used to eliminate some bad matching results.

There are several ways of getting the difference from image matching techniques. As mentioned before, different image matching algorithms can be used in the comparison. Different window sizes in the same image matching algorithm can also used in the comparison to remove bad matching results; or different searching algorithms can be used in the same window size and the image matching algorithm. Different combinations of these can be used for the largest difference between the matching results to get the best matching. In this research only the different image matching algorithms with the same window size and searching algorithm are described.

The combination methods are slower than the single methods, so they suitable for some kinds of work such as automatic relative orientation, determination of the optimal image matching window size for the current stereo pair and for matching some initial points as the first part of the image matching process, and so on.

7.1 Multi-method Image Matching (MMIM) with Whole Range Searching

Methods such as the signal to noise ratio are used to evaluate the matching results. In this section multi-method image matching is described. The so-called multi-method image matching employs more than one image matching algorithm to do the image matching and the matching results are compared, so that the evaluation of the matching results can be done.

7.1.1 Determination of the Threshold Value for the Difference

Assume the CCIM and LDIM and the whole range searching algorithm are used in the multi-method image matching (MMIM). In order to perform MMIM, the threshold for the difference between the matching results from these two matching methods needs to be

defined. There are two ways to do so, one is to define the ground difference, the other is to determine it by statistical analysis.

In the first, assume the difference on the ground is Δh; the heights from CCIM and LDIM for the processed points are h_C, h_L respectively, then if

$$|h_C - h_L| > \Delta h \qquad 7.1$$

for one point, then the matching result for the point is not good. The mean of these two matching results is taken as the final matching result. Figure 7.1 shows the histogram of the difference between LDIM and CCIM for all the 24 check point in the range of window size from 3 by 3 to 95 by 95. Since in Figure 7.1 the differences for most data are 0, only a few matching results have large differences between LDIM and CCIM. So the threshold can be used to remove the matching results with such differences.

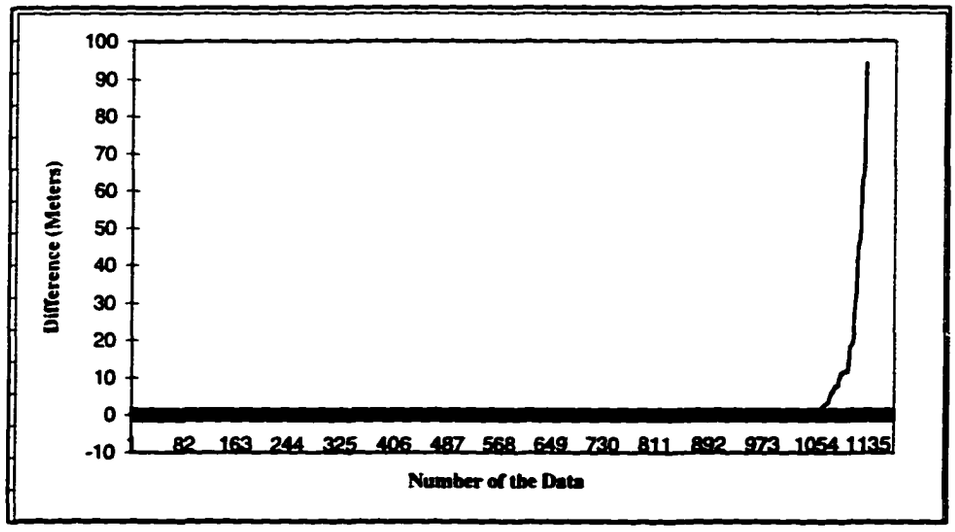


Figure 7.1 Histogram of Difference between LDIM and CCIM

To determine the value of the Δh, the expected height accuracy is considered. For example, if a height accuracy of 0.5 meter is expected, then the value of Δh should be less than 0.5 meter. Since most of the differences are 0, the values of Δh will not greatly affect the multi-method matching results, which means that the value of threshold can be in a relative large range and a few points eliminated.

Using statistical techniques to determine the value of Δh, we assume there are n initial points used and their differences are Δh_i for i = 1 to 20. The standard deviation δ and mean m for the results of these twenty points can be calculated, then define the Δh as

$$\Delta h = m + \delta \qquad 7.2$$

Doing so we can define the matching result in more than two categories such as good, fair and poor. Since the Δh and the matching accuracy are not directly related, the second method may not be accurate.

7.1.2 Matching Differences and the Window Size

Figure 7.2 shows the mean of differences between LDIM and CCIM Vs. the window size for the 24 check points.

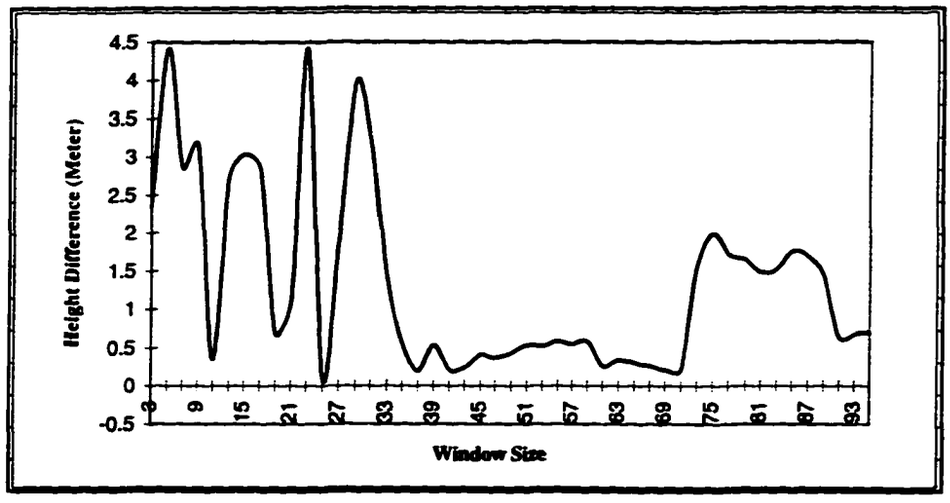


Figure 7.2 Matching Difference between LDIM and CCIM vs. Window Size

In Figure 7.3 the differences are displayed in four categories according to the type of ground feature. All the results shown in the Figures 7.1, 7.2 and 7.3 are the differences between LDIM and CCIM without modification. From Figures 7.2 and 7.3 it is known that the matching differences are large except in the window size of around 37 by 37 and 71 by 71. The matching difference for the ground points is less than that for others. Comparing Figures 7.2, 7.3 and 6.17, matching is good in the similar window size range. The similarity between them shows that the best window size range can also be obtained by analyzing the difference

between the matching results of LDIM and CCIM for some initial points, although the correct heights of these initial points are not known. This conclusion gives one possible method to determine the optimal window size for the image matching.

Figures 7.2 and 7.3 only show the relationship between the matching differences and the window size. They can not tell how much the multi-method improves matching accuracy. This will be discussed in the following sections.

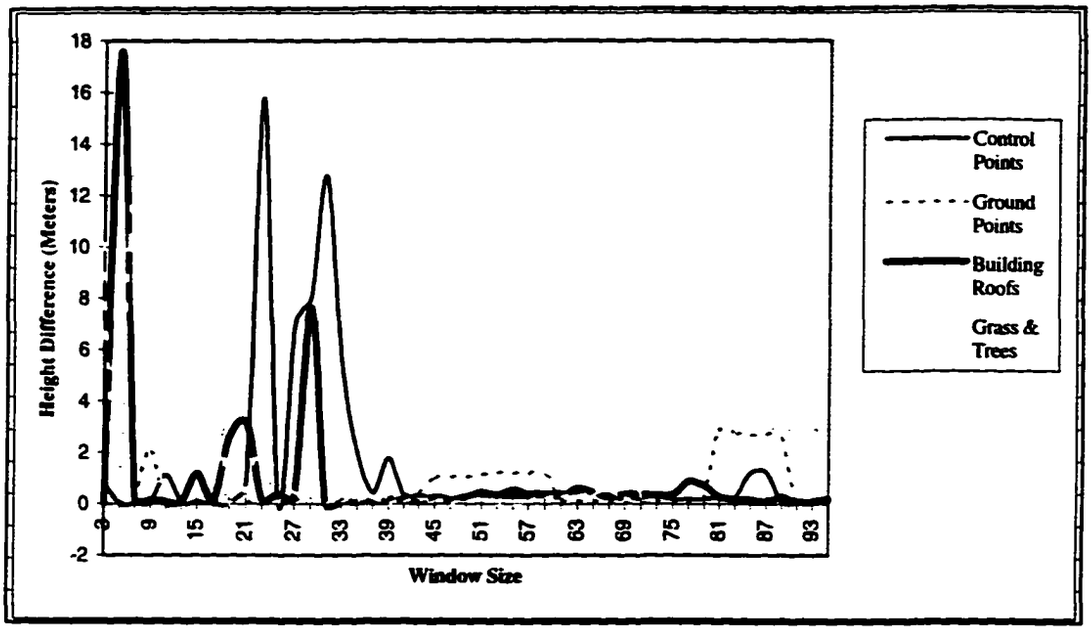


Figure 7.3 Matching Difference between LDIM and CCIM vs. Window Size for the Four Types of Ground Features

7.1.3 Improvement of Accuracy by the Multi-method

The accuracy comparison for Δh = 0.5, 1 and 2 meters is shown in Figure 7.4. From Figure 7.4 it is seen that although the threshold values are different, the accuracy difference between them are small almost throughout the whole range. Figure 7.5 shows that the number of points removed because they exceed the threshold value is the same for these threshold values. Beyond window size 41 by 41, no points exceed the threshold value.

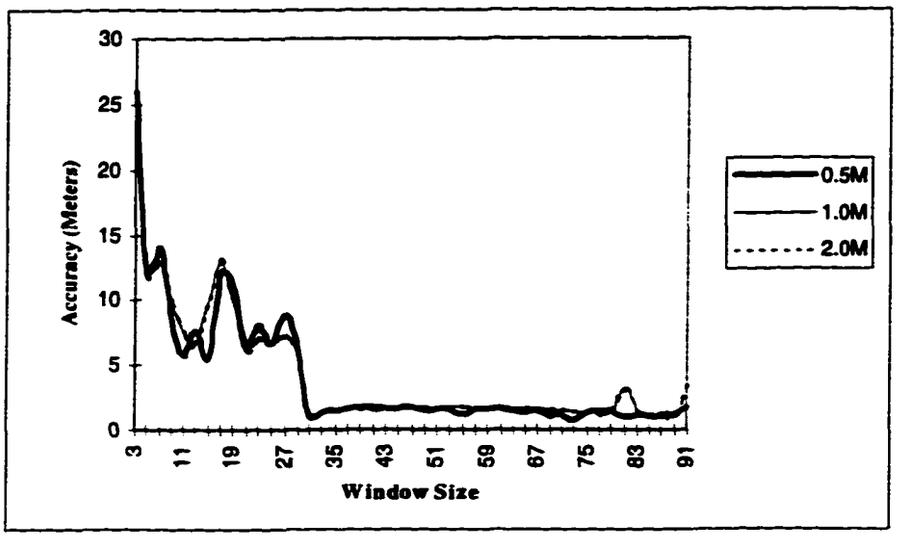


Figure 7.4 Comparison of Accuracy between Different Threshold Values

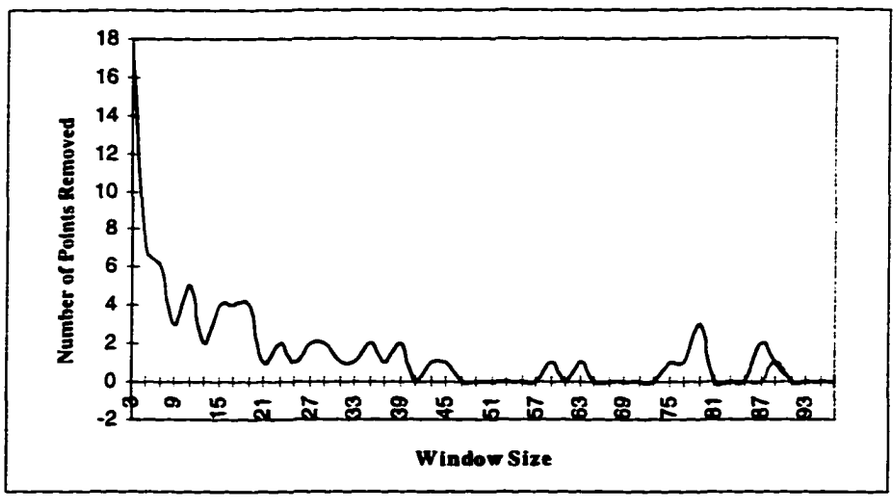


Figure 7.5 Number of Points Removed for Different Threshold Values

Figure 7.6 shows the accuracy comparison between CCIM and MMIM for Δh = 0.5M. When the window size is less than 27 by 27, the MMIM is better than CCIM even by a few meters at some window sizes. The difference after the window size of 75 by 75 is about a few meters. When the window size is between 31 by 31 and 73 by 73, the difference is at a minimum of about half a meter. According to Figure 7.5, the number of points removed diminishes with increase of the window size, and the accuracy of MMIM increases slowly.

The reason for the lower accuracy of CCIM is that a few points have large differences between the matched locations and the actual ones. Therefore MMIM can help to remove some of the points with bad matching results.

In Figure 7.6 the CCIM has higher accuracy than MMIM for some window sizes. This is because the MMIM take the mean of LDIM and CCIM as the final matched location. If the LDIM results are worse than the CCIM, this will happen.

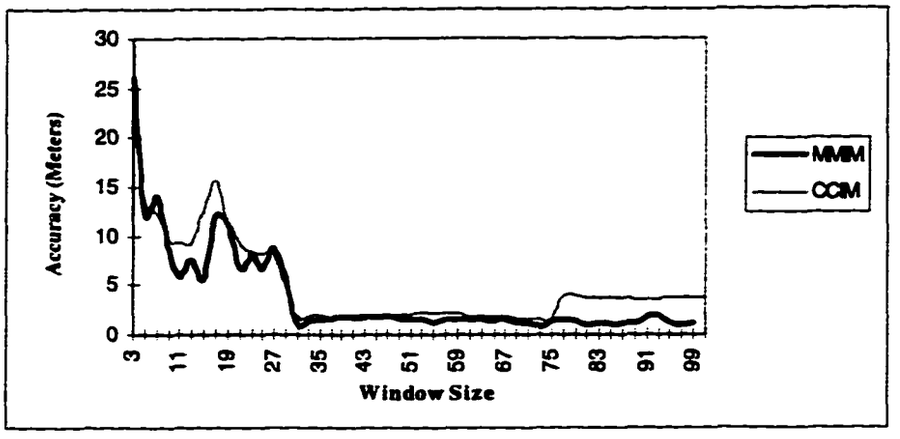


Figure 7.6 Accuracy Comparison between CCIM and MMIM with $\Delta h = 0.5M$

From the above discussion we see that the MMIM can improve the matching by increasing the matching accuracy when the window size is large, and the difference curve is similar to the accuracy curve described in chapter 6 so that it can be used to determine the optimal searching window size. The obvious disadvantage of MMIM is that it is slower than the single method image matching because it employs more than one image matching method.

7.2 Multi-method Image Matching with Best-track Searching

The idea in the multi-method image matching is to make use of the differences from the image matching algorithms so that the errors in the matching can identify the matched position from different matching algorithms. In this way the incorrect points can be detected. The previous chapter shows that the best-track searching algorithms have less constraint since only

the pixels on some rows and columns are used in the matching. So perhaps best-track method can reveal more differences between CCIM and LDIM.

7.2.1 Matching Difference Vs Window Size with Best-track Searching

The comparison between the whole range and best-track searching algorithms is shown in Figure 7.7. According to Figure 7.7, the height difference from the best-track is much larger than from whole range searching, though from their trend that they have the same pattern can be seen. The best-track can have a large difference between the two matching methods in MMIM. It will give better results if these differences reflect the impact from the errors on the image matching algorithms and not from the mismatching.

Figure 7.8 shows the comparison between whole range and best-track using MMIM for Δh=0.5 M. It is obvious that the best-track has a more unstable accuracy when the window size is less than 35 by 35. But the difference between these two methods is very small when the window size is larger than 35 by 35. The matching results from best-track is better than the results from whole range searching when the window size is large. So the unstable matching results from the small window size is because of the mismatching. When the window size is big enough, the mismatching is less, and the large difference from the best-track gives better matching results than whole range.

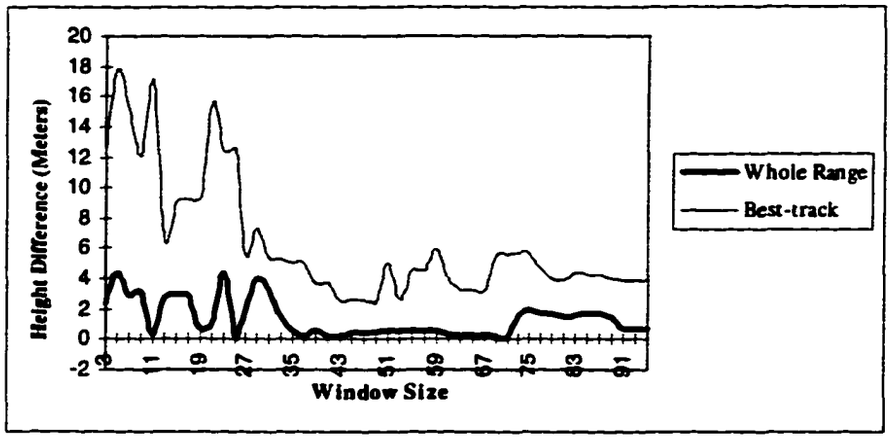


Figure 7.7 Comparison between Whole Range and Best-track Searching for MMIM

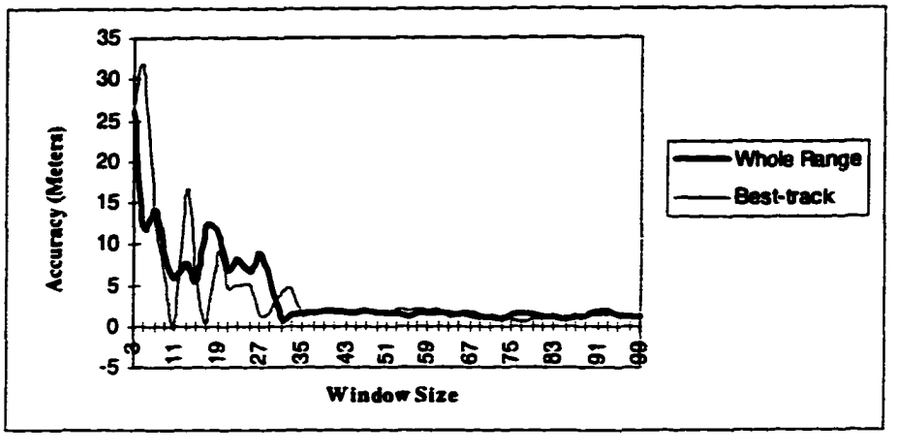


Figure 7.8 Accuracy Comparison between Whole Range and Best-track with MMIM for Δh = 0.5M

As shown in Figure 7.9, the best-track has a almost linear relationship between time consumed and the window size. The relationship between speed and window size for whole range is that the square of the time consumed is proportional to window size. So the larger the window size, the slower the whole range searching. Since the best-track searching is much faster than the whole range searching, especially for large window size, it could be used in the matching of the initial points, and for the whole stereo pair.

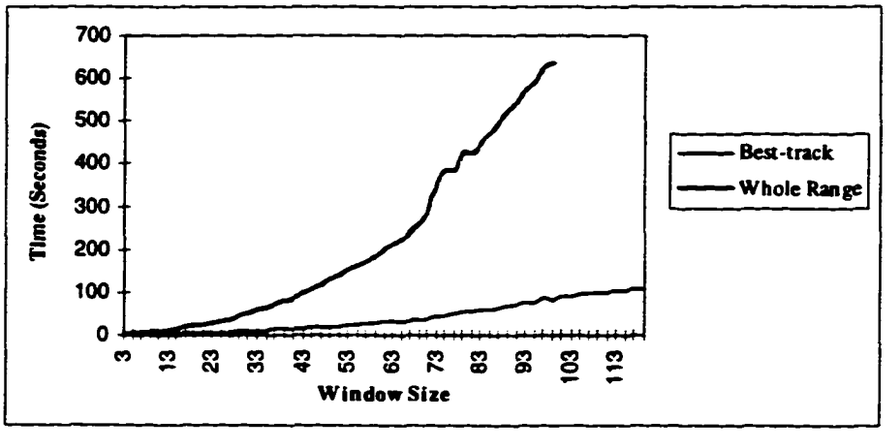


Figure 7.9 Speed Comparison between Whole Range and Best-track Searching

7.3 Determination of Optimal Matching Window Size

Since area-based image matching is performed on a number of pixels in the certain area defined by the searching window, the matching process is affected by the quality of the photographs used in the matching, types of the ground features, scale of the photograph, matching algorithms, searching strategies such as the searching range in both directions, and the window size. Apparently, the window size determines the information which can be used in the matching. For a certain area of a photograph, there should be an optimal window size for the best matching results, for a stereo pair, there should be a window size for the best matching in terms of statistics. Multiple window sizes for a stereo pair can be considered if the selected area contains different types of ground features with different characteristics. In this case, the division of the area is the first step that should be done. Only one optimal window size is described here.

7.3.1 Window Size and the Signal to Noise Ratio (SNR)

Since window size actually contains the information used to do the image matching, the Signal to Noise Ratio (SNR) determines the amount of correct information in the matching. The information in a window consists of (1) background noise; (2) correlated information; and (3) uncorrelated information. A good matching requires (a) enough correlated information and (b) enough SNR. The first means that the correlated information in the window on each photograph in the stereo pair needs to be large enough for the matching; the second means that the ratio of correlated information to the uncorrelated information and the background noise also needs to be sufficiently large. The background noise is from any noise that could appear in the data processing such as in film exposing and developing, imagery scanning and so on. The correlated information is given by the similarity of the pixels in the windows during matching. The uncorrelated information could come from different sources such as ground relief and wrong window position. The coefficient can not be one even for both windows on the same spot on each photograph in the stereo pair because of the change of brightness between the exposures, and the geometric and obstruction changes caused by the different exposure positions. So the uncorrelated information and the background noise can be considered together as matching noise. The SNR is the ratio of correlated information to the

matching noise. So the optimal window is the combination of larger SNR and proper processing speed.

In terms of correlation coefficient the SNR can be written as:

$$SNR^2 = \frac{r_c^2}{1 - r_c^2}$$

for the cross correlation image matching. Therefore the larger the r_c, the larger the SNR. Normally smaller window can have higher r_c because of less matching noise and more similarity as there are fewer elements in a smaller window. Figure 7.10 shows the relationship between the window size and the value of r_c. The coefficient r_c drops from almost 1 to 0.97 when the window size is about 113 by 113.

So the larger the window size, the smaller the SNR. But SNR only gives the evaluation on the fitness of the correlation between the two windows at a certain window size, which does not necessarily mean that the larger window will give the worse matching results. Only the SNRs of similar window sizes can be compared.

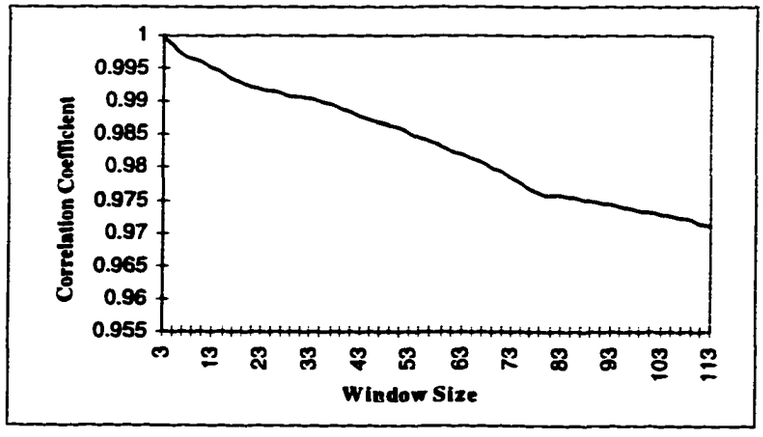


Figure 7.10 Correlation Coefficient vs. Window Size

In order to get the best matching result, the largest ratio between correlated and uncorrelated information is needed. The assumption here is that, at first, the correlated information increases faster than the uncorrelated information as the window size increases when the window size is small, the ratio increases until to its maximum with the best matching

results. After the window size with the maximum ratio, the uncorrelated information increased faster than the correlated information and the matching worsens. The relationship between the window size and the mean of the height differences between the matching results and the correct heights for these check points is shown in Figure 7.11 using the cross correlate image matching method. In this stereo model, the mean of the height difference for the smallest window of 3 by 3 is about 18 meters, then it drops to about 4 meters at the window between 7 by 7 and 9 by 9. It has the smallest height differences at the window sizes between 31 by 31 and 77 by 77, thereafter the difference increases to about 4 meters.

From the previous discussion it can be concluded that from window 3 by 3 to 7 by 7, the correlation increases; from window 7 by 7 to 17 by 17, the uncorrelated information and noise increase faster than the correlated information; from window size of 13 by 13 to 31 by 31, the correlated information increases faster than the uncorrelated information, so the height difference is decreasing; the increase of all the information remains balanced from window 31 by 31 to 77 by 77, then uncorrelated information is increasing faster than the correlated information and the accuracy is decreasing. From the Figure 7.11 it can be concluded that after a certain window size, the image matching gets stabilized. This conclusion is based on all the 24 check points.

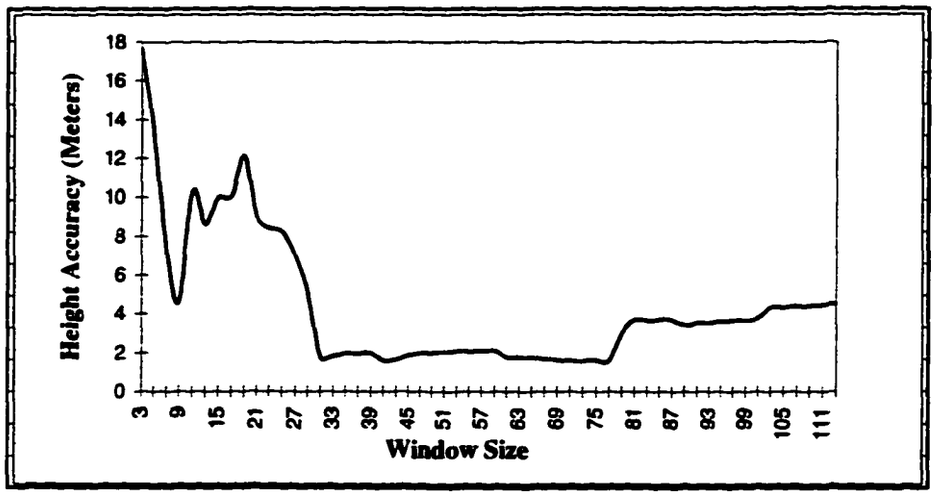


Figure 7.11 Mean of Height Differences vs. Window Size

At each point, the balance between these types of information will be different depending on the type of the ground feature.

7.3.2 Use MMIM to Determine the Optimal Matching Window Size

The curve shown in Figure 7.11 is not known in real cases. Therefore it is not possible to use it to determine the optimal window size. From the above analysis we know that the difference between LDIM and CCIM is at a minimum when the window is the optimal size. Therefore the difference between LDIM and CCIM can be used to detect the optimal window size. The similarity between the height difference curve from MMIM and the accuracy curve in Chapter 6 has been mentioned.

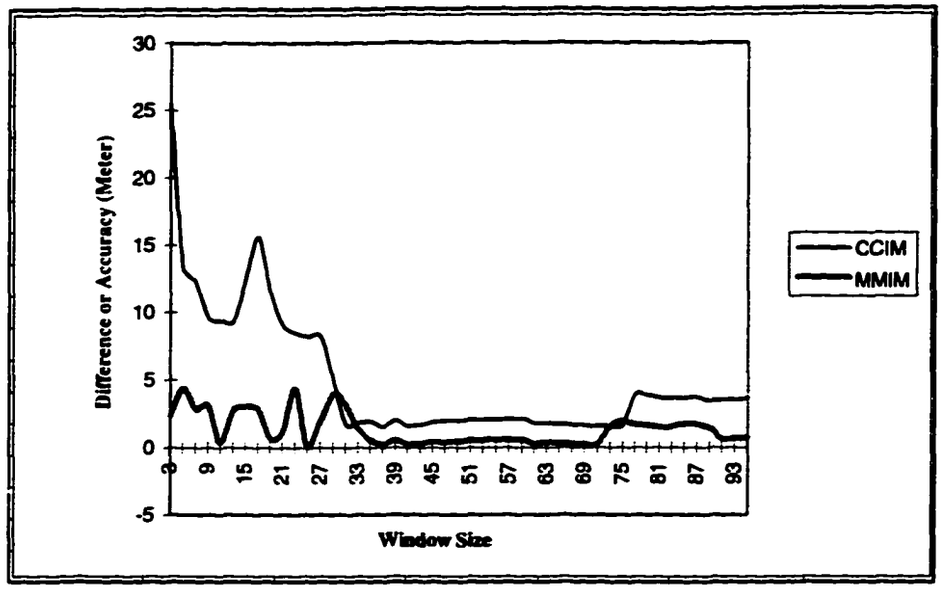


Figure 7.12 Comparison between Difference from MMIM and Accuracy from CCIM

As shown in Figure 7.9, both curves have the stable low values in the same area. In other words, they have the same pattern. It seems that the low area of MMIM curve is covered by the low area of CCIM curve. This finding can be used to determine optimal window size for image matching. Note that in Figure 7.12 the curve of CCIM is accuracy vs. window size, the curve of MMIM is the difference between CCIM and LDIM vs. window size.

The procedure used to determine the optimal window size is described by the flowchart in Figure 7.13. Basically, it looks for the smallest matching window size with a low and stable difference between CCIM and LDIM in MMIM. Since the difference shown in Figure 7.12 for MMIM could be small before reaching the range of optimal window size, we must ensure that the low difference is stable. Before determining the optimal matching window size, a certain number of initial points should be selected on the stereo pair to be matched.

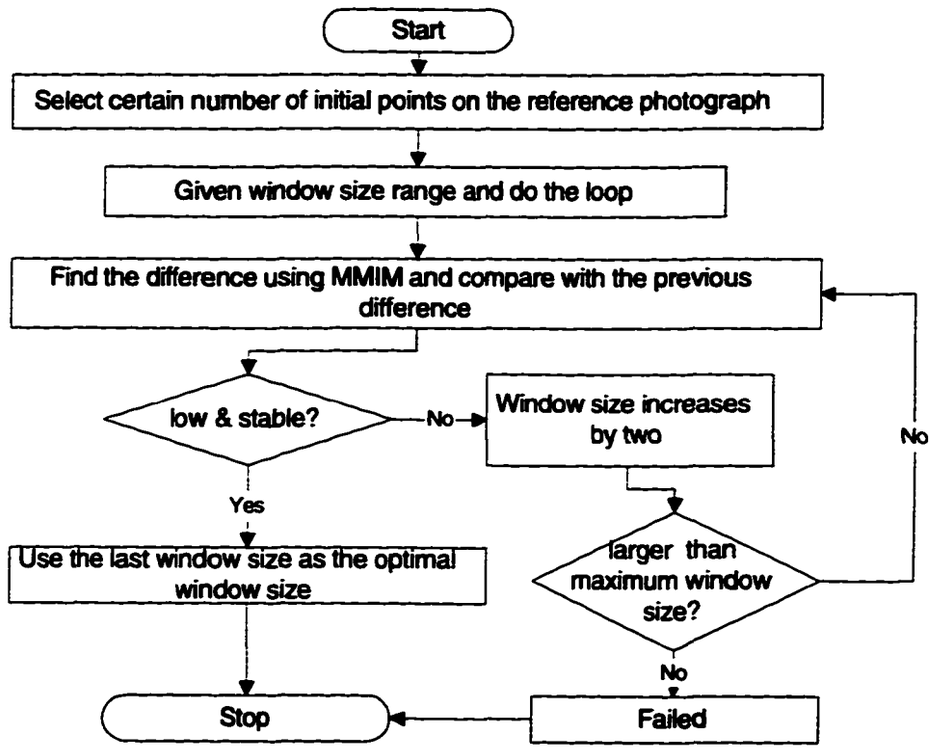


Figure 7.13 Flowchart of Using MMIM to Determine Optimal Window Size

7.3.3 Use Number of Points Agreed in MMIM to Determine the Optimal Window Size

Assume that the matching has the most correct position when the SNR is the largest. So there are the most agreed positions when SNR reaches the maximum because of the uniqueness of correct positions. Therefore, the window size of maximum SNR can be determined by comparing the matching result between consecutive window sizes. As the

window size increases, the SNR may be decreasing because of more similarity than the dissimilarity; therefore the number of agreed points may be increasing. The optimal window size is detected if the matching from two consecutive windows agree the most. It is possible that there are many optimal window size at which the matching is good, and only the smallest is used. Therefore, to search for smallest optimal window size is to search for the first maximum number of agreed points. The flow chart of this process is shown in Figure 7.14. Any area-based image matching algorithms can be used in this process.

Since the ratio of similarity to dissimilarity varies when the window size changes from small to large, a false maximum could appear, as shown in Figure 7.16. To remove the error of the false maximum, a number of window sizes will be checked after the first maximum is reached. If the previous maximum is exceeded at a larger window, then this number is taken as the new maximum and the same procedure is repeated until no new maximum is found. The number of window that needs to be checked depends on the number of initial points and the images. For the case shown in Figure 7.16 eight more windows are sufficient. Therefore, we used 10.

To distinguish the MMIM for consecutive windows from the MMIM discussed before, we call the former double MMIM, and the earlier single MMIM. Double MMIM means there are two threshold values used in the process of matching, one to check the matching difference at the same window size when doing MMIM, another one to check the matching difference between the consecutive window sizes.

Figure 7.15 shows the relationship between the window size and the number of positions matched using double MMIM and the whole range searching algorithms. Figure 7.16 shows the results from best-track. Although they do not have a similar curve pattern, their first maximum occurs at almost the same window size. So the best-track searching can be used instead of the whole range searching because of its faster speed. Since the best-track searching has a large difference between the two image matching algorithms in MMIM than that with whole range searching method, the number of agreed points in Figure 7.15 is more than in Figure 7.16.

If the process shown in Figure 7.14 is used to search the optimal window size then, from Figures 7.15 and 7.16, the optimal window size will be about 33 by 33 which matches the optimal window size obtained in the previous discussion.

As shown in Figure 7.17, double MMIM can improve the matching accuracy when the window size is small, in this case, less than window size of 29 by 29. When the window size is larger, the difference between them is not significant although the result from MMIM is still less than that from CCIM.

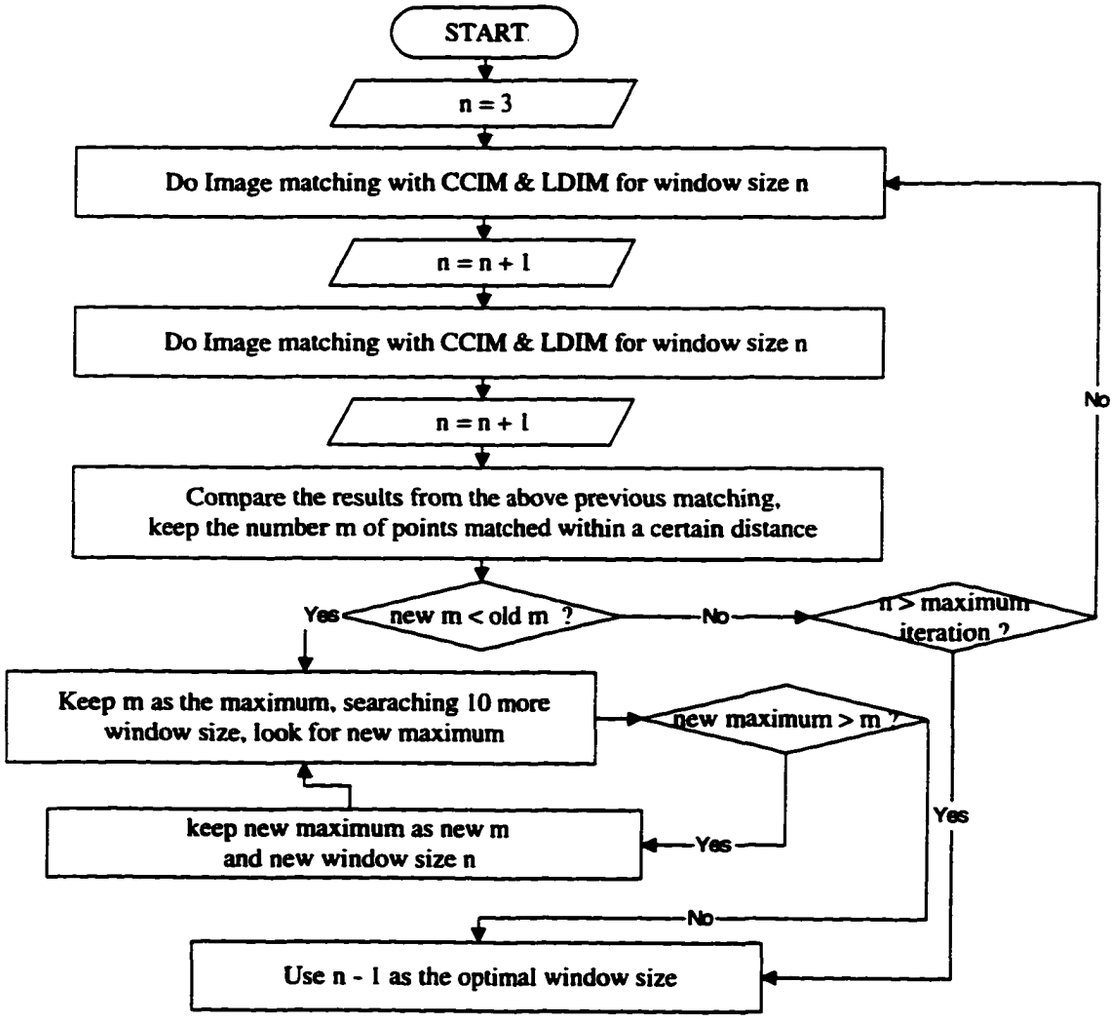


Figure 7.14 Flow Chart for Determing the optimal Window size Using Double MMIM

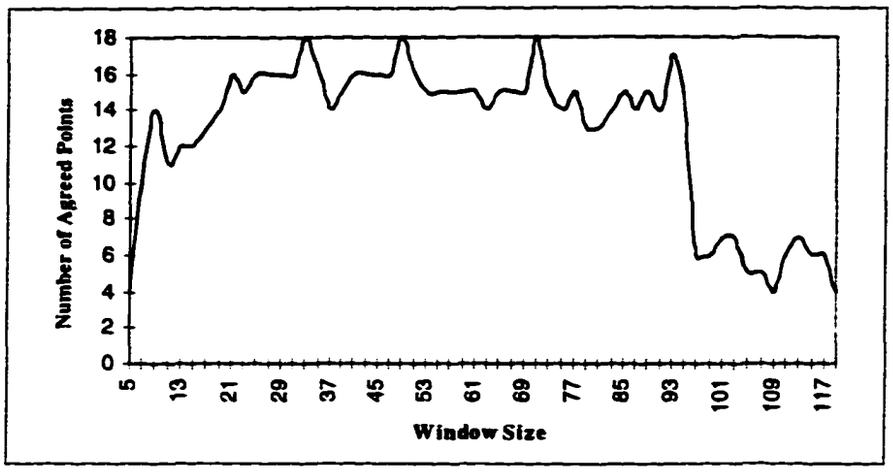


Figure 7.15 Number of Point Agreed in Double MMIM with Whole Range Searching

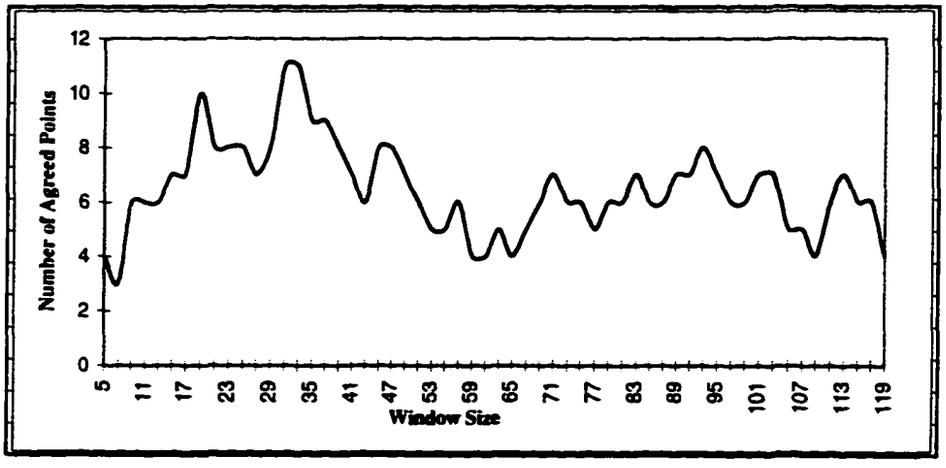


Figure 7.16 Number of Agreed Points in Double MMIM with Best-track Searching

7.4 Test on Random Selected Points

The above conclusions are the matching results from the 24 check points. In this section a number of randomly selected points are used to test these conclusions.

Since the 24 selected check points have relatively good location or contrast which the randomly selected initial points of the real situation may not have, it is necessary to test them on the randomly selected initial points. These initial points are evenly distributed over the whole reference image.

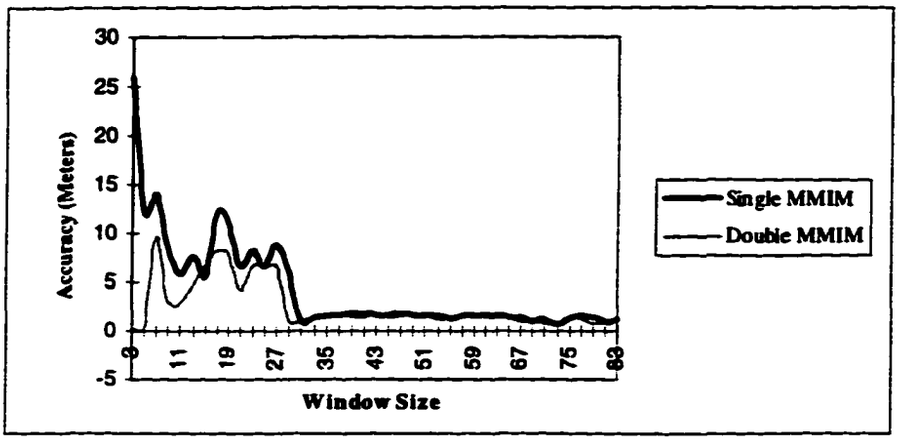


Figure 7.17 Comparison between Single and Double MMIM

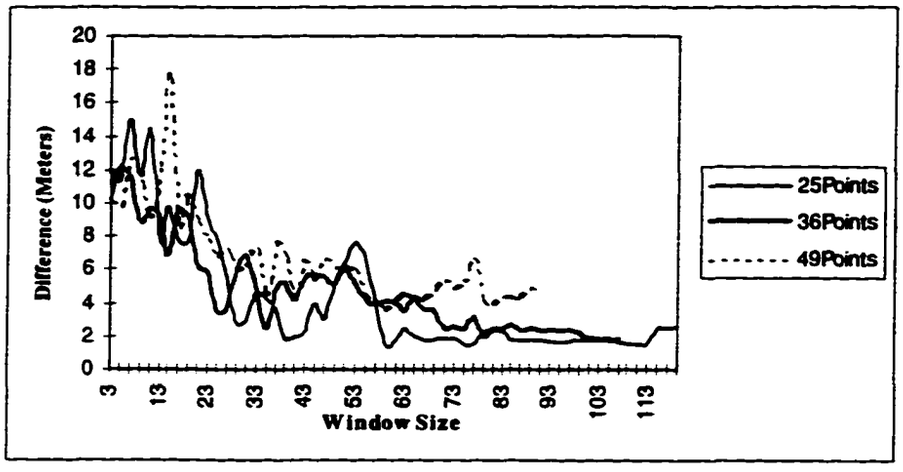


Figure 7.18 Difference Curve for Different Number of Initial Points

7.4.1 Difference Curve for Different Number of Initial Points

Figure 7.18 shows the curves for differences Vs. window size for 25, 36 and 49 initial points. Since the best-track searching is used these curves are not very stable, but they all have the same pattern. The optimal window size using varying initial points is different, but they are all in the range of good matching results. According to Figure 7.7, the matching difference from the best-track searching is not as stable as that from the whole range searching, so it may be better if the whole range searching is used.

Another problem here is how to determine the optimal window size using the curves in Figure 7.18. Unlike the curves in Figure 7.7 in which the optimal window size is easy to select. Some statistical method may be needed to determine the optimal window size.

7.4.2 Number of Agreed Points for Different Number of Initial Points

Figure 7.19 shows the number of agreed points Vs window size for different number of initial points for different threshold values.

The first two lines in Figure 7.19 shows the comparison between Δ=0.5M and Δ=1.0M for 25 initial points. It is understandable that the curve from Δ=1.0M is above that from Δ=0.5M because the threshold value for the latter is larger. But the optimal window sizes from both are almost the same, which is around 41 by 41. 16 initial points are also tried but no result is available. From a statistical point of view, the similar results can be obtained if the number of initial points is sufficient.

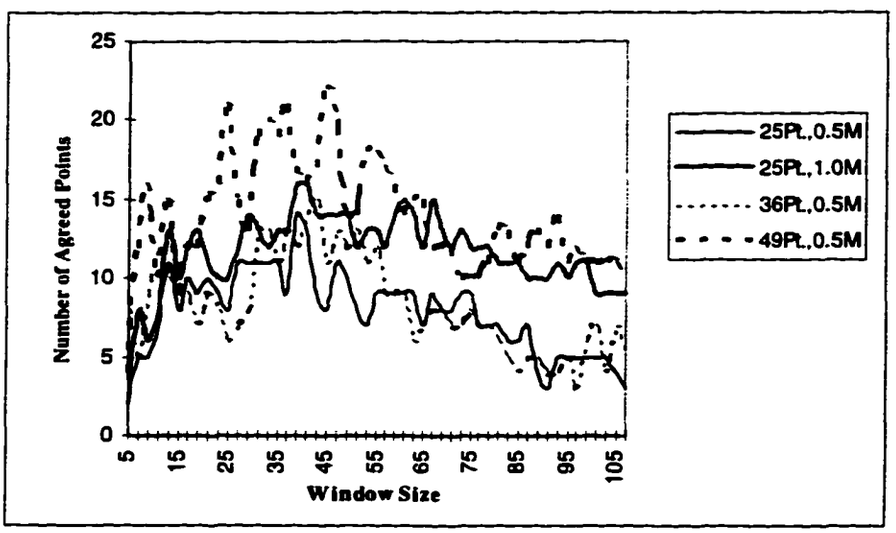


Figure 7.19 Number of Agreed Points vs. Window Size for Different Number of Initial Points

According to Figure 7.19, the optimal window size obtained from different initial points vary but are very close, in the range 41 by 41 to 47 by 47.

According to the above results, the optimal window sizes from those randomly selected initial points are larger than those from the 24 check point. Statistically the number of agreed

point decreases after a certain window size since the noise or dissimilarities will be more than the similarities after that window size. Unlike the 24 check points, some of the randomly selected initial points may have low contrasts, so a larger window might be needed to get the optimal matching results. The optimal window size is larger for the randomly selected initial point, which is closer to the real situation.

7.5 Conclusions

From above discussions it is known that the MMIM can be used to determine the optimal window size according to the difference between the image matching algorithms in the MMIM. A different number of initial points can be used in the MMIM process. This is an automatic process which does not receive a priori knowledge of the height control points.

CHAPTER 8 UNSQUARE WINDOW IMAGE MATCHING

The windows used in the previous chapters were squares with equal number of rows and columns. As mentioned in Chapter 5, the window with more rows than columns should be better than the square window. In this chapter the comparison between them will be presented.

8.1 Unsquare Window Image Matching

Assume the pixel is a square, then in an unsquare window the number of rows is not equal to the number of columns. After epipolar rectification, the x axis is along the flight direction. As discussed before, for the photographs in the same stereo pair the x parallax is much larger than the y parallax because of the larger difference in distance in the flight direction. So unlike in the case of x direction, the viewing angle in y direction is almost the same for every point, and the obstruction for its image on the photographs in the stereo pair should be almost the same. The conclusion here is that there are more similarities in the y direction than in the x direction. The window should extend more in the direction of y axis, and the window with more rows will give more useful signals than the square window does when the number of pixels are the same.

In the following discussion, m will be standing for the number of columns of the window, n will be the number of rows of the window. It is a square window if $m = n$.

When $m > n$ we shall refer to it as x window. When $m < n$ it is called y window. For the same 24 check points the image matching results will be compared.

Unlike in square window, the number of pixels and not the window size is used here as the x axis of the figures in the comparison.

8.2 Comparison between Different Types of Unsquare Windows

8.2.1 Comparison between the Mean Values with Fixed Number of Rows and Columns

One way to analyze the unsquare window is to get matching result by keeping one dimension and increasing the other. The matching results with fixed number of rows and fixed number of columns are compared.

Figure 8.1 shows the comparison between the results when the number of columns is fixed.

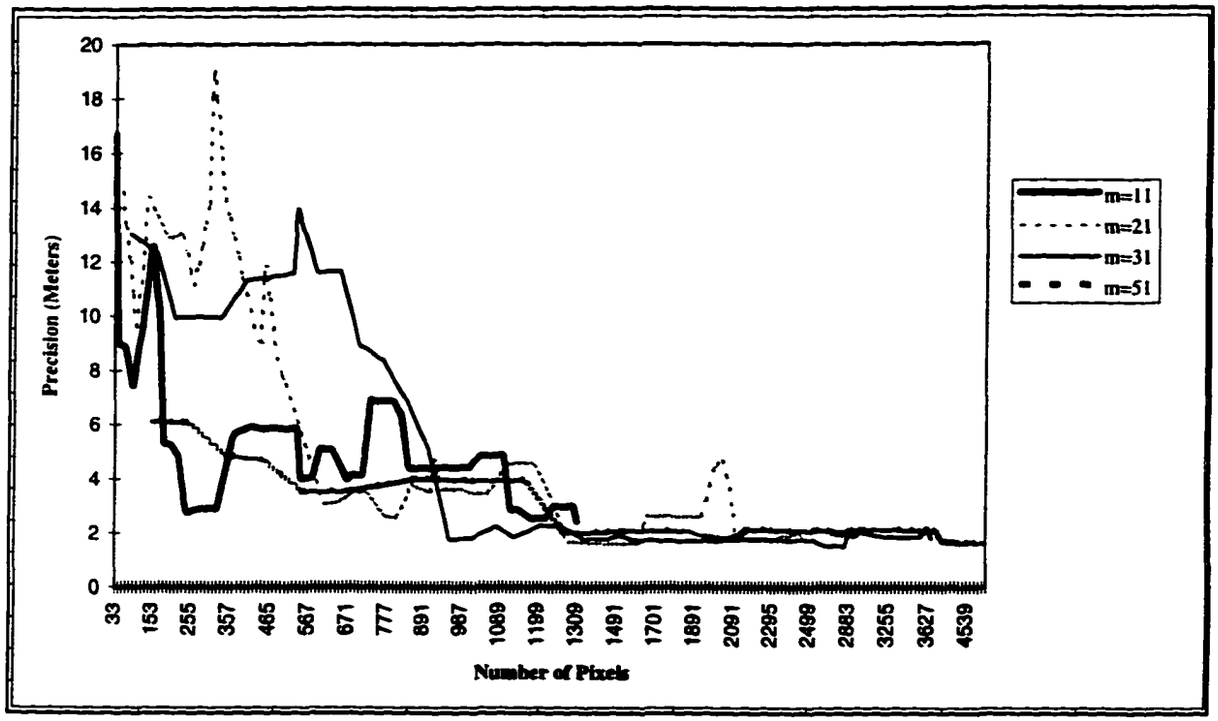
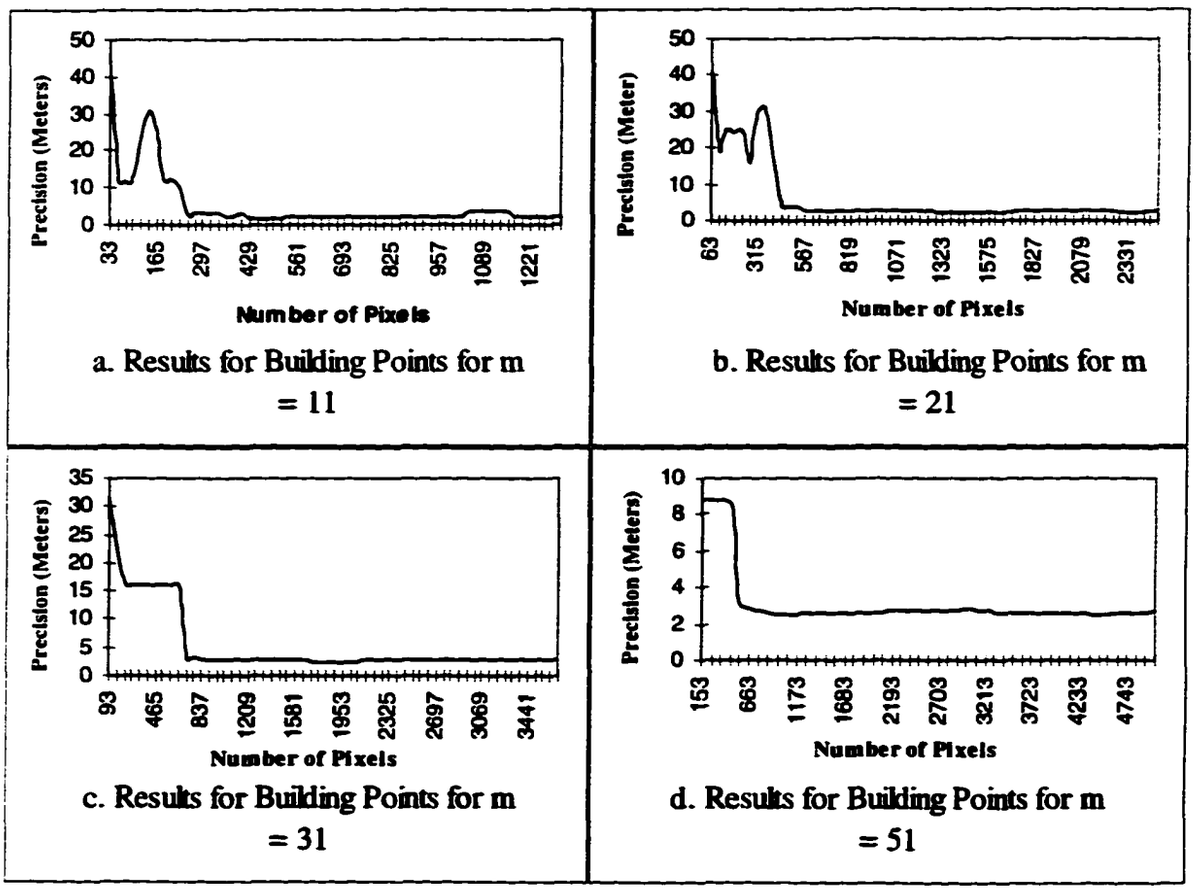


Figure 8.1 Results for Given m

The window with m = 51 has the best precision of about 6 meters when the number of pixel is less than about 200. After 200, with m = 11 precision increases rapidly to about 2.5 meters. All the curves have almost the same precision when the number of pixel in the window is large. It is obvious that the curve with m = 51 is the most stable, although it has the lowest precision when the pixel number is larger than about 2300. The above figure shows that as the

number of rows or n increases and the number of columns or m remains constant, the matching precision improves.

Since the roof points have the most height variance in the matching window, the results for the building roof points for different m values are shown in Figures 8.2 a, b, c and d. The curves for m = 11 and 21 decreases when the window size in the y direction is about 15 pixels. For m = 31, the curve decreases when the window size in y direction is about 26 pixels. For m = 51 the window size in y direction is about 12 pixels when the curve begins to decrease. The curve for m = 51 looks more stable than others, but it has a little larger errors than the others.



a. Results for Building Points for m = 11

b. Results for Building Points for m = 21

c. Results for Building Points for m = 31

d. Results for Building Points for m = 51

Figure 8.2 Results for Building Roof Points for Different Values of m

Figure 8.3 shows the comparison between the results when the number of rows is fixed. Compared to the results shown in Figure 8.1 it has much lower matching precision. In the beginning all the curves except $n = 11$ drop fast. This is because that the number of rows for $n = 11$ is too small, the information for matching from the y direction is not sufficient. the curve for $n = 11$ starts to drop at the pixel number of about 500. Since the number of rows is fixed, and does not increase with the window and the number of columns, the information ratio between x and y will decrease as the number of pixels increase. Because of this, the best precision for $n = 31$ and 51 is about 2 meters; and for $n = 21$ and 11 is about 4 meters.

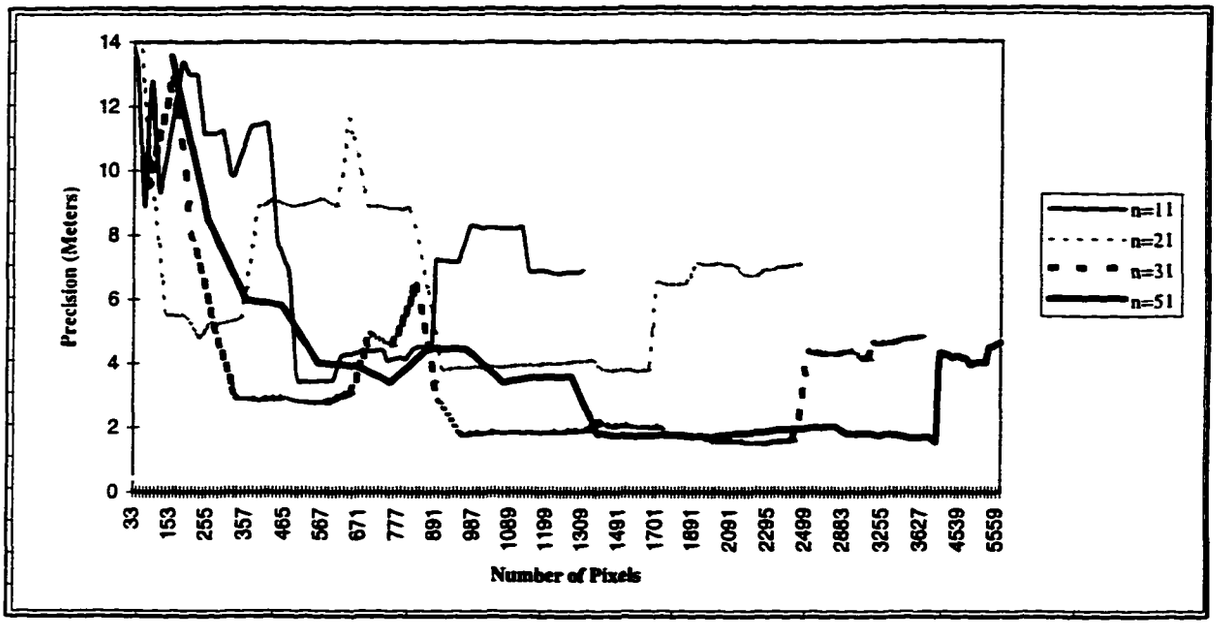


Figure 8.3 Results with n Fixed

Figure 8.3 shows worse results than that shown in Figure 8.1. In Figure 8.3, all curves rise sharply when the number of columns is about 80. The main reason for this is the sharply elevated roof points, they have the worst results through the whole range of pixels. The ground control points do not have good matching results until there are about 1400 pixels in the window. This is because the roof points have the most ground height change and so they have the most dissimilarities in the x direction. When the window size in x direction is larger

than a certain number, the matching precision will drop due to the sharply increasing dissimilarities. Since these ground points are selected as corner points or intersection with good contrast, they have good matching results. The control points are targeted points on the ground which do not necessarily have good contrast.

Figure 8.4 shows the precision between the four different types of ground features for n = 51.

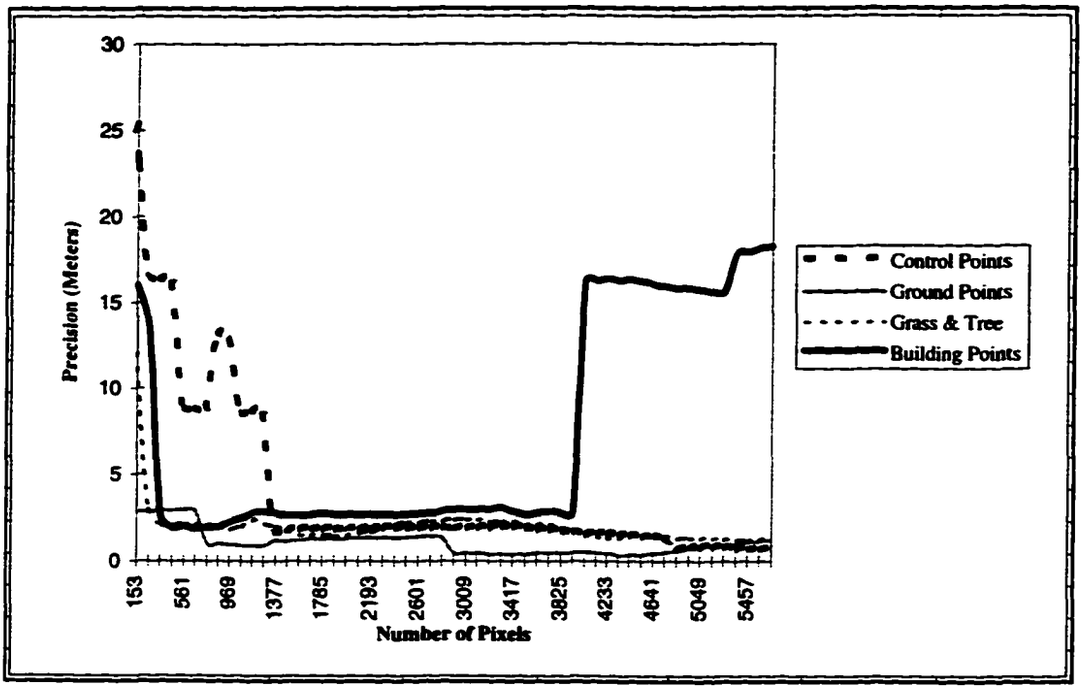


Figure 8.4 Comparison between the Four Types of Ground Points for Number of Row n = 51

Figure 8.5 shows the results for the four types of ground points for n = 11. Compared to Figure 8.4, only the ground points still have good matching results. In fact, the ground points have better matching results in Figure 8.5 than it is in Figure 8.4. Because in the feature the ground points have good contrast in a small area. As the matching window increases, the contrast could decrease. The curve for roof points also rises sharply when the number of columns of the matching window is about 80 pixels. Even considering the different scales of the x axis, the ground points have the best results shown in Figure 8.5.

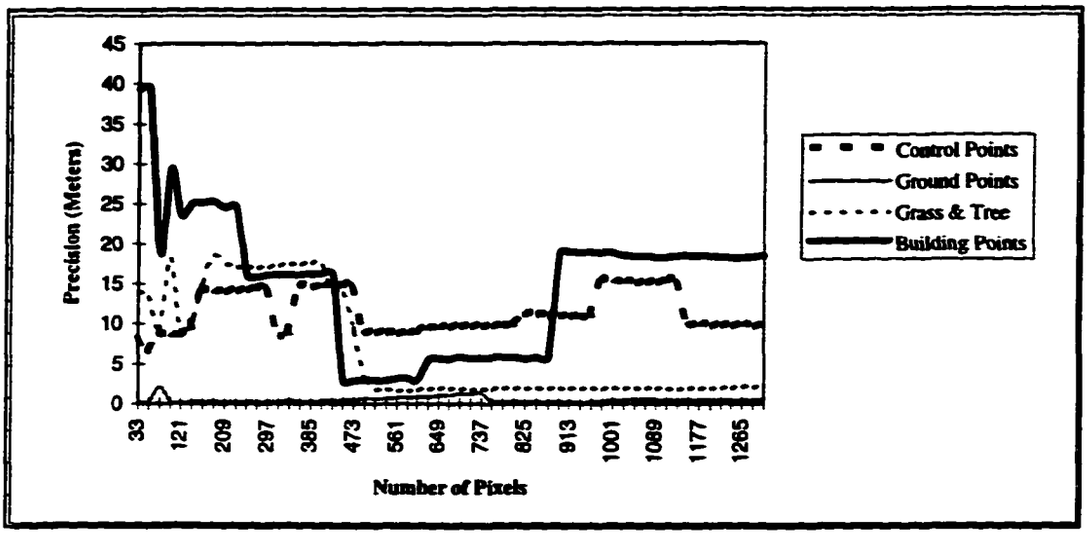
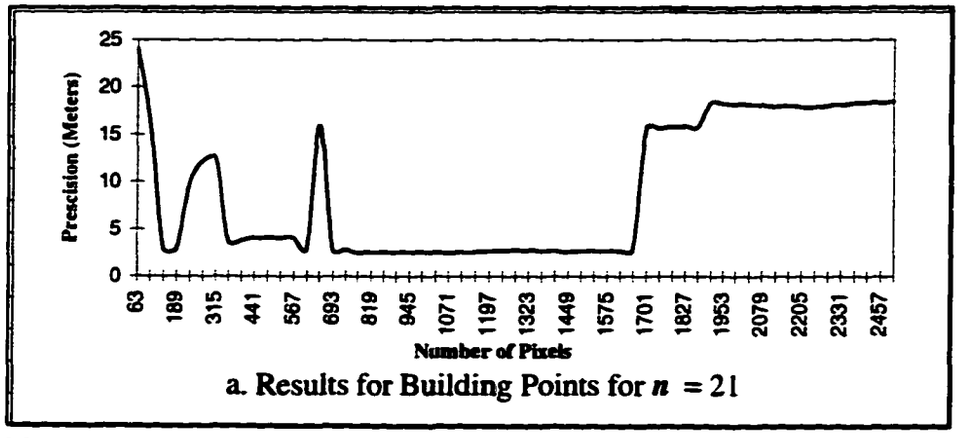
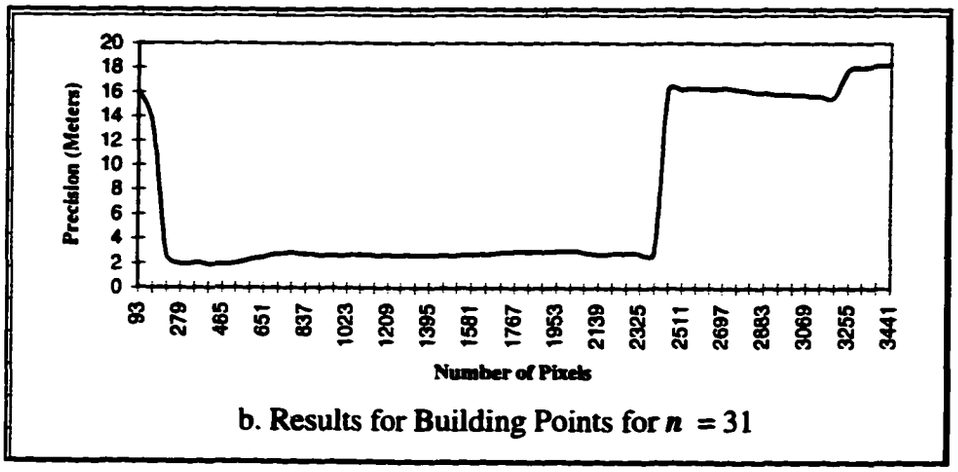


Figure 8.5 Comparison between the Four Types of Ground Points for Number of Row n = 11



a. Results for Building Points for n = 21



b. Results for Building Points for n = 31

Figure 8.6 Results for Building Roof Points for Different Values of n

Figures 8.6 a and b show the results for roof points for $n = 21$ and $n = 31$. Comparing the curves for building roof points in Figures 8.4, 8.5, 8.6 it can be seen that the best results are when n is the largest number, 51; and the worst results are when n is the smallest number, 11. Also comparing these with the results shown in Figures 8.2a to 8.2d, these results have much worse matching precision. Therefore the larger window size in y direction can to a certain extent that is when the number of columns are below a certain number improve matching of roof points.

The above comparison can be done for other types of points.

From the above discussions and figures the conclusion is that the larger number of n can give a better matching results for an area with large height change.

8.2.2 Comparison between the Check Points with Certain Shape of Window

The above section discusses the results from windows with fixed m or n. In this section the results from the window size of a certain ratio between the two dimensions of the window are discussed.

Figure 8.7 shows the comparison for two windows with $m = 2*n$ and $2*m = n$.

The curves in Figure 8.7 can explain some of the differences between x and y windows. When the number of pixels involved in the matching is less than 1000, the matching with the y window has much better results than that with the x window, though the difference decreases. Therefore the pixels along y axis are more informative than pixels along x axis, the same conclusion from the previous section. This is more obvious when the ground height changes dramatically in the searching area due to the presence of buildings which will be shown later. Since different types of ground points are used in the 24 check points, the curves in Figure 8.7 are only the mean values of them. On the other hand, for x windows, since the number of rows increases as the number of columns increases, the information from pixel along y axis also increases. Therefore, the difference between x and y windows could decrease and be close together a certain level. In Figure 8.7 this happens when number of pixels ranges between 1400 and 3000. After that, the difference between these two curves increases again with the increasing window size. This means that the increases in x direction bring more noise than useful signals making the matching precision to drop for the x windows.

Both sides of the window are increasing simultaneously according to the given ratio; therefore, regardless of what this ratio is, the window in *y* direction will increase with the increasing number of pixels. The results will be different from those shown in the previous section. Even the size of the window is larger in *x* direction than in *y* direction.

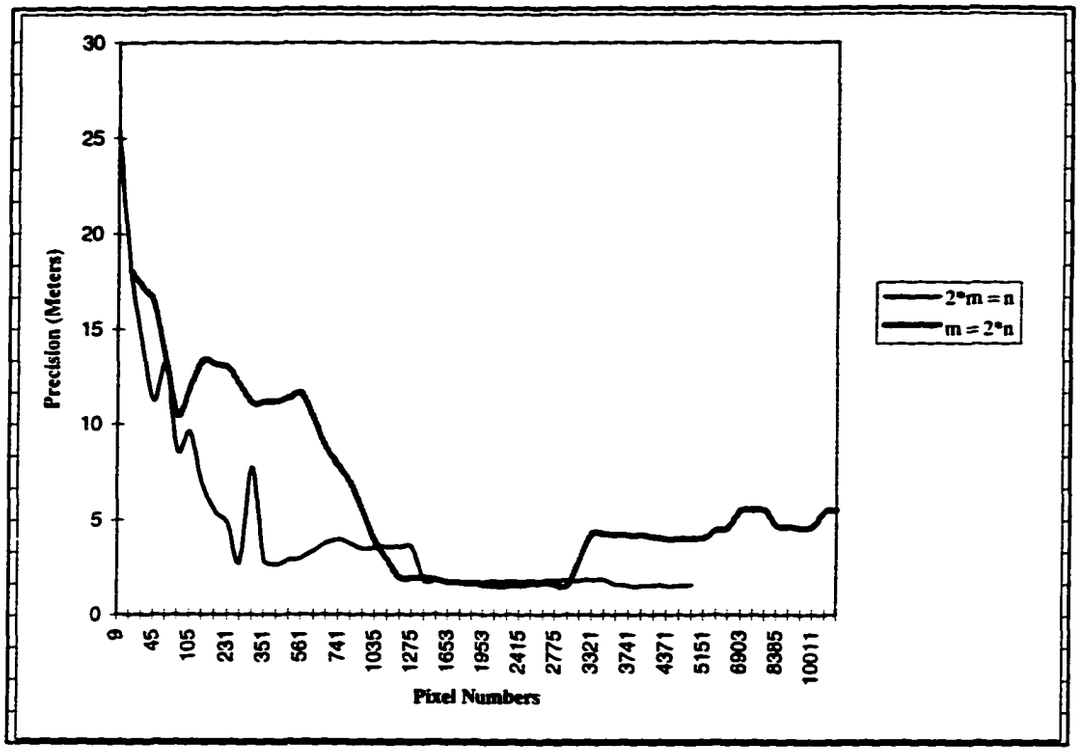


Figure 8.7 Precision Comparison between Window Shape $2*m = n$ and $m = 2*n$

Figures 8.8 a, b, c, d show the comparison between the different types of points.

The thinner line in each figure represents the *y* window with the larger number of rows. It can be seen that, in general, the *y* window has better results than the *x* window for all four types of ground points. The differences for grass and tree points and roof points are larger than that for control points and ground points because of the height variance in the matching window for roof and the tree points. For the control points, the *y* window is better than the *x* window when the number of pixels is less than about 1000, after where they have close results.

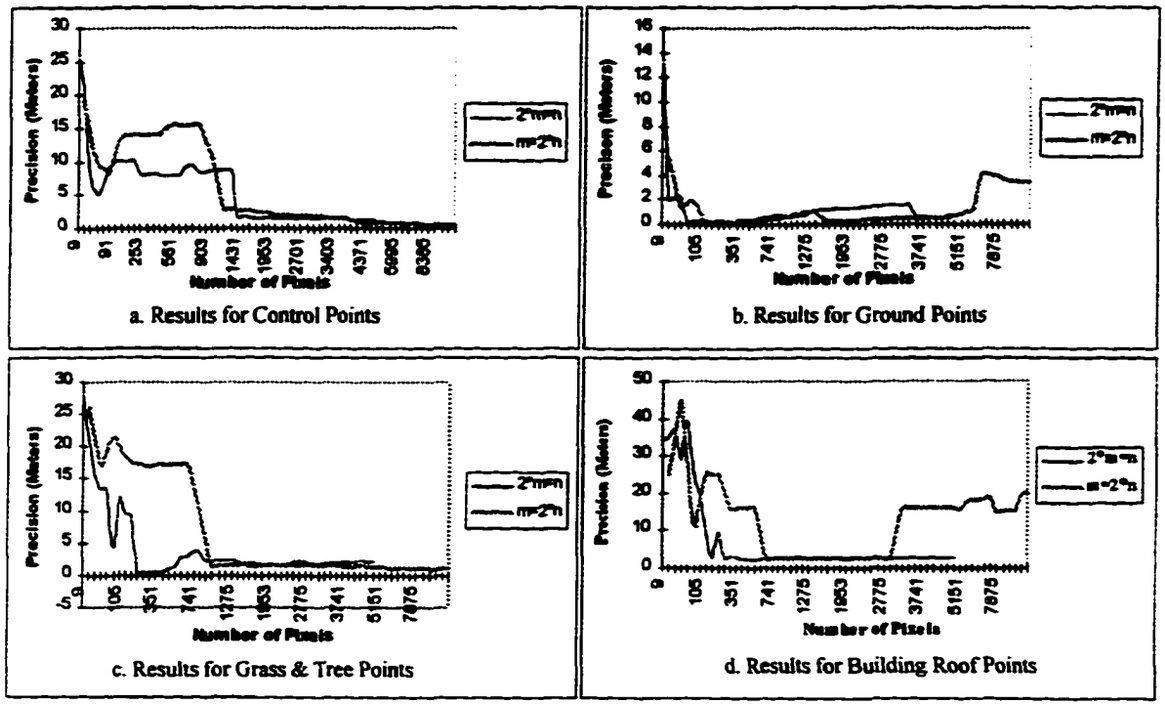


Figure 8.8 Comparison between Results for Different Types of Points

For the ground points, the *y* window is slightly better than *x* window when the number of pixels involved is less than about 150. The results are close when the number of pixels is between 150 and 1270. The *x* window has even better results when the number of pixels is between 1270 and 3700, after which the curve from *x* window rises sharply. Since the ground points have good contrast and location, the shape of the window does not really matter. If there are obstructions in the *y* direction, the results from *y* window will be worse than that from the *x* window.

For the grass and tree points, the *y* window has better results when the number of pixels is less than 1000, after which the results are close.

For the roof points the *y* window has better results when the number of pixels is between 120 and 740 or larger than 2800. They has close results when the number of pixels is between 740 and 2800.

8.2.3 Comparison between Different Shape of *y* Windows

The comparison in this section is about the *y* windows with a different ratio between the number of rows and the number of columns. Figures 8.9 a, b, c, d, e and f show the comparisons. Notice the different scale in these figures.

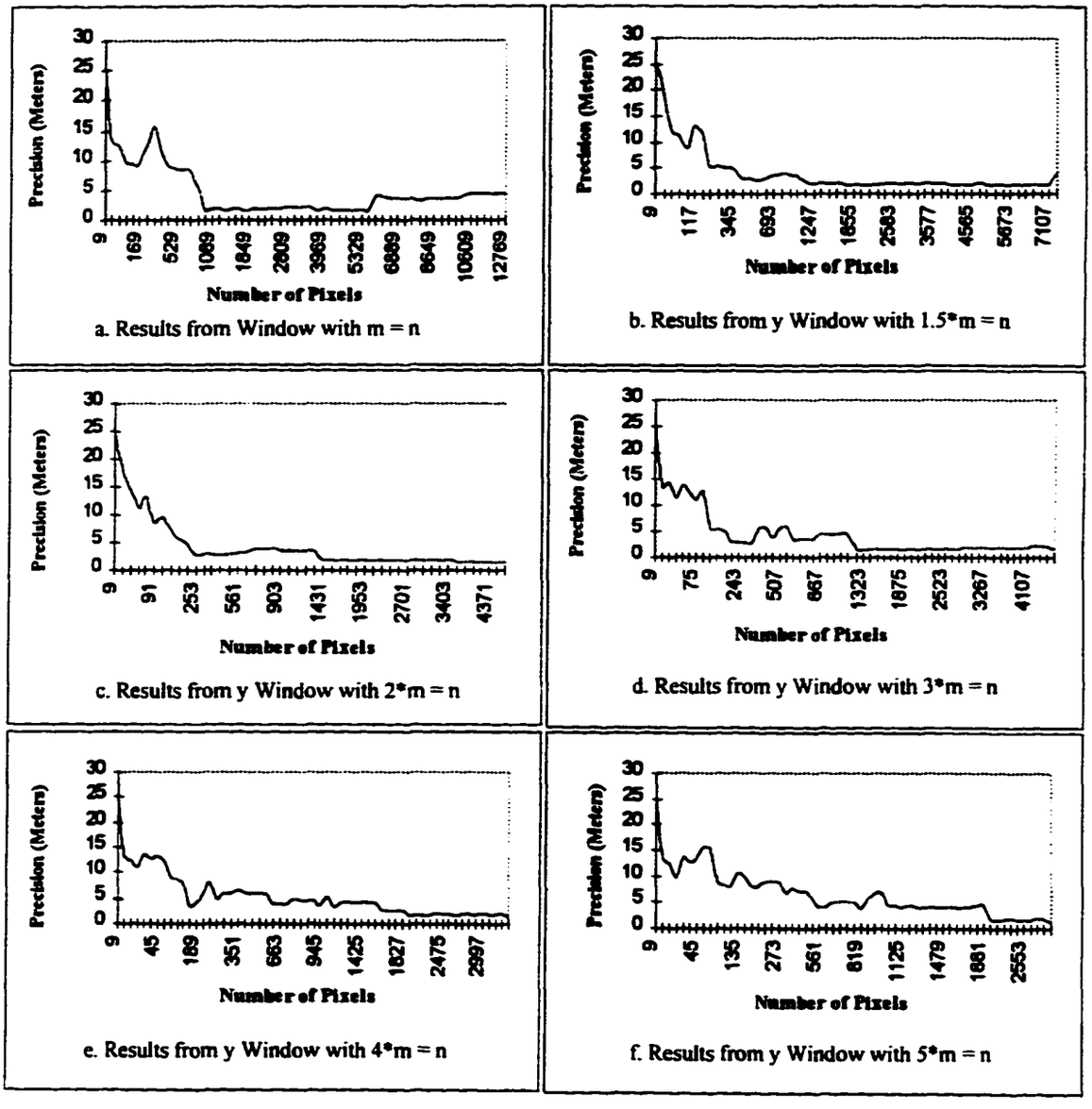


Figure 8.9 Comparison between Different *y* Windows

Assume $k*m = n$ for $k \geq 1$. The number of pixels at which the curve yields good results increases with the increasing value of k from 1200 for k = 1.5 to about 2160 for k = 5. But the larger the k, the better the best results for a certain range of k values. Accordint to Figure 8.10, a $k \approx 3$ is the best.

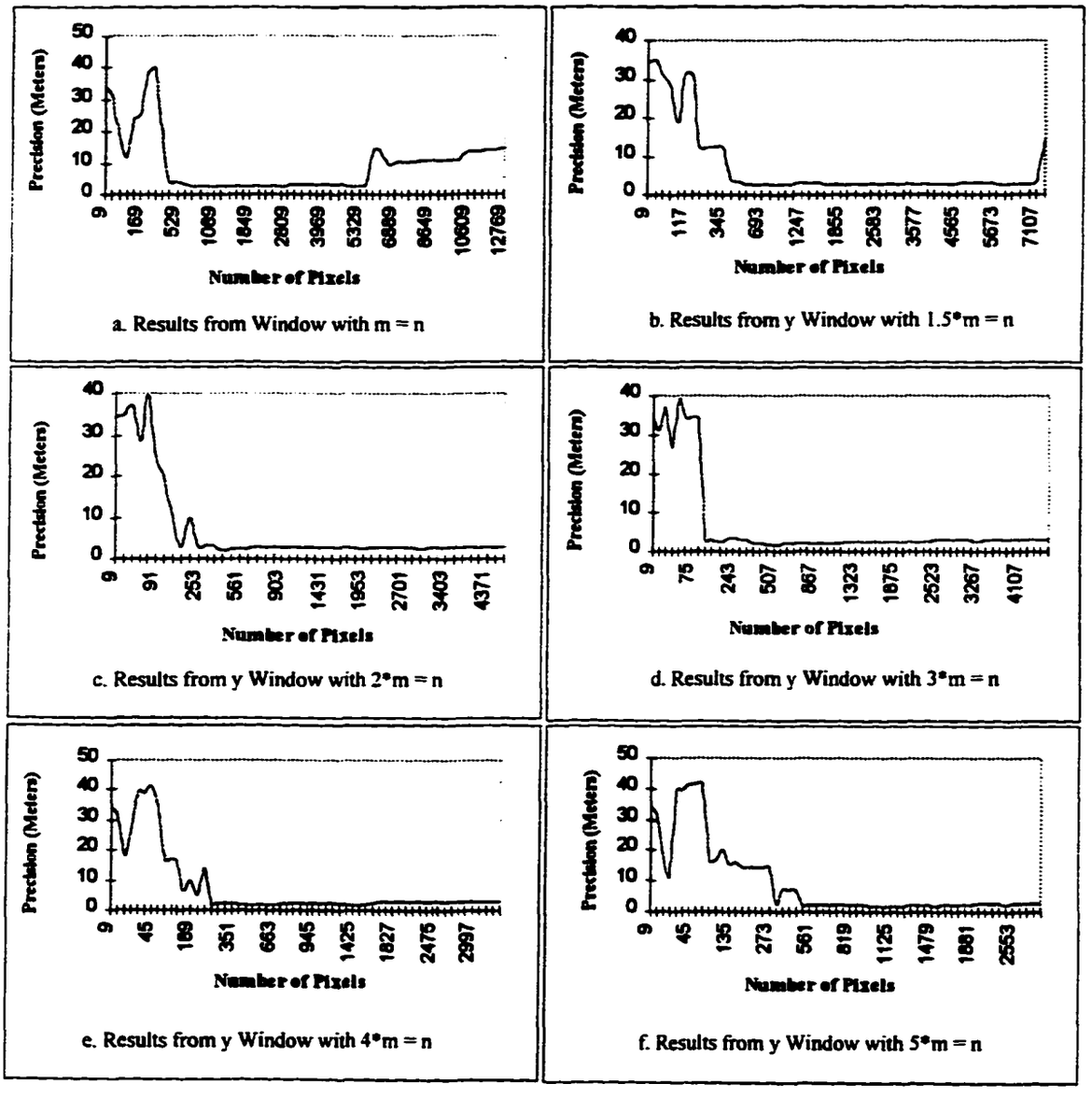


Figure 8.10 Comparison between Results from Different *y* Windows for Building Roof Points

The comparison between *y* window and the square window is also shown in Figure 8.9 a. Although the curve from square window drops earlier than the *y* window, it increases again when number of pixels is about 5500, while the curves from *y* window do not increase; and it has worse matching results than that from the *y* windows even in the good matching area. Therefore, the *y* window has better matching results.

The results for the roof points are used here as example for the different types of ground points since they have the largest ground height change.

Figures 8.10 a, b, c, d, e and f show the results for roof points for different shapes of windows. From these figures it can be seen that the larger the **k**, the better the results. So a larger **k** or getting more information from the *y* direction can improve the matching results for the roof points.

Since the square windows are also included in the above figures, the comparison between the square and unsquare windows can also be seen.

8.3 Conclusions

The above results from the comparison between unsquare window and square window indicates that the unsquare window can give better matching results, especially when the *y* window. The conclusions are:

- A larger window in the *y* direction can have better matching results than the larger square and *x* window, especially for the area with extreme height changes.
- More information from the *y* direction is helpful improving the image matching.
- Fewer pixels may be needed in getting the best matching results.
- The image size will limit the size of the unsquare window.

CHAPTER 9 CONCLUSIONS

In this thesis, research was carried out on image processing, image matching searching strategies, feature type and image matching, and optimal window size. A comparison was also made between CCIM, LDIM and LSIM. Two new image matching methods, multi-method image matching and unsquare window image matching were also developed. In the process of doing so, several issues in soft photogrammetry were discussed. The conclusions and recommendation are presented in this chapter.

9.1 Image Pre-processing

Image pre-processing processes the raw images so that they are in stereo pairs and can be used in image matching. The steps involved in the pre-processing are image orientation and rectification. The format used to store the processed images and output DEMs need to be considered. In this research, a modified LAN image format has been developed to store all the information necessary in image matching.

9.1.1 Modified LAN Image Format

The standard LAN image format is for *Landsat* imagery. A modified LAN format is used in this research to store the information on image matching. Chapter 2 presented the modified LAN image format. The advantage of the modified LAN image format is that fewer files are needed in a project. The images in the modified LAN format can still be used by software like IMAGINE which recognizes the standard LAN format.

The size of the normal LAN header is 128 bits. Some of which are not used. For now the unused bits are enough for the purpose of this research. If the information needed is more than the unused space, more bits can be used.

9.1.2 Interior Orientation and Rectification of Photographs

Interior orientation and rectification make the photographs ready for image matching. Image matching will be better if the photographs in the same stereo pair have the same scale and pixel size, and are in epipolar coordinates.

In order to obtain the same pixel size, a resampling process that makes all the photographs the same pixel size is used in the stereo rectification. Also, a process that makes the photographs in the stereo pair the same scale is applied in the stereo rectification. Therefore, the stereo rectification not only transforms the photographs into an epipolar coordinate system, it also makes them the same scale and pixel size.

As discussed in section 4.3.8, an error will be introduced in the equalizing scale process. This error can be removed by finding the correct position after image matching. Therefore, the above resampling process will not introduce any additional principal errors.

For better resampling results, different types of interpolation methods are applied in the process of orientation and rectification.

All the results from rectification were checked with SoftPlotter and the comparisons are good.

9.1.3 Direct Transformation

By using direct transformation, all the steps discussed in the previous sections are done in one step. Before the direct transformation can be done, the information about the interior and rectification should be known. Thus, the direct transformation is a combination of the equations used in orientation and rectification. The difference between this two types of processing is that at least two photographs need to be used in the direct transformation because of the stereo rectification. In the previous processing, only one photograph is needed until the stereo rectification is done.

The advantage of direct transformation is that only one resampling is applied in the whole process. Therefore, the impact on the images from resampling is kept at a minimum.

9.1.4 Equations and Their Programming Implementation

The equations suitable for programming implementation are derived and shown in chapters 7, 8. The results obtained using these equations compare with those obtained by SoftPlotter confirming the correctness of these equations. These include the determination of the parameters used in the equations and for the images. The comparison with SoftPlotter shows their correctness.

The programs used in the process are for measuring fiducial points, interior orientation, single photograph rectification, stereo rectification, equalizing pixel size and photograph scale.

9.2 Image Matching Searching Strategy

Two types of searching strategies in image matching are discussed in the same searching range. One is called whole range searching in which all the pixels within the searching range will be used in image matching. Its description and discussion are presented in the section 6.4.1. The other is called best-track searching in which only the pixels in the searching path are involved in image matching. It is described and discussed in the section 6.4.2. The comparison between these two types of searching strategies is also presented.

9.2.1 Comparison between Best-track and Whole Range Image Searching Strategies

The comparison between them are performed in terms of speed and matching precision. The best-track searching is much faster than the whole range searching since fewer pixels are involved in image matching. For matching precision, there is almost no difference between them when the window is big enough. The conclusion is that the best-track searching is better than the whole range searching. Therefore, application of the best-track searching can save time in image matching.

Since four different types of ground points in four different types of ground features are used in this process, the conclusion should be good for most of the cases.

9.3 Study on Feature Type and Image Matching

The correlation coefficient surface from CCIM and the absolute difference surface from LDIM for the 24 check points are presented in chapter 6. According to the figures showing the surfaces, it can be concluded whether the matching results are good. This is the graphical way to analyze the relationship between the type of ground feature and the quality of image matching algorithm. The conclusions are:

1. Points with good contrast and surrounding continuity will give good peaks on the surface graphs and good matching results.
2. CCIM is more sensitive to surrounding discontinuity

3. LDIM is more sensitive to low contrast

9.4 Study on Determination of Optimal Window Size

The optimal window size for different stereo pairs and different types of imageries could be different. It is related to a number of factors such as the type of imagery and digitizing device (scanner), the brightness and contrast used in scanning, the pixel size and geometric precision of the digital imageries, and the characteristics of ground features in the image matching. Therefore, the determination of the optimal window size for a stereo pair is a dynamic process. The optimal window size could be different in different parts of the stereo pair. The multi-method image matching is used to determine the optimal window size.

9.4.1 Single MMIM in Determination of the Optimal Window Size

Single MMIM method uses the similarity between the difference curve of MMIM and the precision curve of CCIM shown in Figure 7.12. Both curves have similar low value ranges. The low value range from CCIM corresponds to the range of the good image matching. So the low range of the difference from MMIM is used to determine the optimal window size.

The stability of the low difference from MMIM is needed to obtain the correct optimal window size.

9.4.2 Double MMIM in Determination of the Optimal Window Size

In double MMIM the window size increases when image matching for the check points are performed. The number of points within the height difference from the image matching algorithms is compared with the number of good points from the previous window size and the maximum number is recorded. The window size with the maximum number of good matching is considered as the optimal window size. The results on the 24 check points show that this method gives the optimal window size which is consistent with the previous results on the relationship between the precision and the window size.

9.5 Comparison between CCIM, LDIM and LSIM

The LSIM is the fastest in all three methods when 4 iterations are done. Since LSIM is sensitive to the initial position, and it needs to use CCIM or LDIM to get the initial position, it may not be the fastest method.

The results show that LSIM combined with CCIM has higher precision than CCIM alone. The results from the roof points also show that LSIM is sensitive to surrounding discontinuity.

9.6 Study on Image Matching Algorithms

Two new image matching methods, multi-method image matching and unsquare window image matching are discussed. They are described in chapters 7 and 8.

9.6.1 Multi-method Image Matching (MMIM)

Any image matching algorithm should give the perfect results when the images in the matching are identical. The dissimilarities around the image pixels of the same ground points make the results from the matching algorithms to shift sway from their correct matching positions. Generally, the more the dissimilarity that exists, the more the shift. If the same dissimilarity affects different image matching algorithms differently, the shifts will occur in different directions. Therefore, more than one image matching algorithm can be used and their matching results are compared. The matching results can be considered correct if the height difference between the matching results are less than a certain value.

The results in chapter 7 shows that the MMIM can improve the matching results, especially for some window size ranges. The test results also indicates that the MMIM with the best-track searching method is much faster than the MMIM with the whole range searching method.

9.6.2 Unsquare Window Image Matching

Different shapes of windows are checked and compared in the chapter 8. The assumption is that, in image matching, the pixels in the y direction give more useful information than the pixels in the x direction. Here y refers to the columns, and x refers to the rows of the image. The test results support this assumption. The conclusions are:

1. Windows longer in the *y* direction have better matching results than windows longer in *x* direction or square windows.
2. More information in the *y* direction is helpful for improving the image matching.
3. The ratio between the *y* and *x* directions of the unsquare window should be about 2 to 3 for the better matching results.
4. Fewer pixels than in the square window may be necessary to get good matching results.

9.7 Recommendations

The more complicated the algorithm used in image matching, the more time it takes. For a fixed image matching algorithm, a digital hardware correlator can provide a way to solve the conflict between the complexity of algorithm and the speed.

Optical information processing and the combination of optical and electronic system may be one of the ways to obtain high speed image matching. The optical correlator has been used in the automatic target seeking system.

In addition to rectangular window, irregular-shape window can also be used in image matching. The difference between the image areas around the same spot on photographs in a stereo pair could be caused by the moving objects in the area, or some other reasons. Irregular window could be a way to remove this difference. Research can be carried out on how to determine the shape of irregular window in image matching for each point matched.

More study can be done on the determination of optimal window size using double MMIM described in Chapter 7. A more effective and accurate method should be used to determine the threshold value which can be used to determine the optimal window size.

Combinations of rectangular, irregular windows and MMIM can be tested to find a more accurate and faster method which can be used in relative orientation, and determination of optimal window size.

The purpose of image matching is to create DEM from aerial photographs by a certain ground spacing distance. The image algorithms described in Chapters 7 and 8, and recommended in this chapter can be used to create a DEM from the stereo pair used in this study by a certain ground spacing distance.

REFERENCES

Abramson B. S., 1993, "Evaluation of Edge-preserving Smoothing Filters for Digital Image Mapping' ISPRS, Vol. 48, No. 2, pp. 2 - 17.

Ackermann F., 1984, "Digital Image Correlation: Performance and Potential Application in Photogrammety", Photogrammetric Record, 11 (64), pp. 429 - 439.

Ackermann F., 1983, "High Precision Digital Image Correlation", Proceeding 39th Photogrammetric Week, pp. 231 - 243.

Ackermann F., Schneider W., 1986, "High Precision Aerial Triangulation with Point Transfer by Digital Image Correlation", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 18 - 27.

Ackermann F., Schneider W., 1986, "Empirical Investigation into the Precision of Digital Image Correlation", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 115 - 130.

Adamos C., Faig W., 1992, "Hough Transform in Digital Photogrammetry", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 250 - 254.

Al-Tahir R., Schenk T., 1992, "On the Interpolation Problem of Automated Surface Reconstruction", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 227 - 232.

Atkinson K. B., 1980, "Developments in Close Range Photogrammetry - 1", Applied Science Publishers Ltd., London.

Armstrong M. P., Marciano R., 1994, "Inverse-Distance-Weighted Spatial Interpolation Using Parallel Supercomputers", Photogrammetric Engineering and Remote Sensing, Vol. 60, No. 9, Sept. 1994, pp. 1097 - 1103.

Balce A. E., 1986, "Determination of Optimum Sampling Interval in Grid Sampling of Digital Elevation Models for Large-Scale Application", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 40 - 54.

Barbalata I. C., 1992, "Digital Analysis of Error Ellipsoids in Analytical Photogrammetric Stereotriangulation", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 639 - 726.

Barnard S. T., Thompson W. B., 1980, "Disparity Analysis of Images", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 4, pp. 333 - 340.

Benard M., Boutaleb A. K., Kölbl O., Penis C., 1986, "Automatic Stereophotogrammetry: Implementation and Comparison of Classical Correlation Methods and Dynamic Programming Based Techniques", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 131 - 140.

Benard M., 1984, "Automatic Stereophotogrammetry: A Method Based on Feature Detection and Dynamic Programming", Photogrammetria, Vol. 39, pp. 169 - 181.

Bertram S., 1963, "The Automatic Map Compilation System", Photogrammetric Engineering, Vol. 29, No. 4, pp. 675 - 679.

Bethel J. S., Mikhail E. M., 1984, "Terrain Surface Approximation and On-line Quality Assessment", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 25, Part A 3a, Commission III, pp. 23 - 32.

Blachut T. J., Burkhardt R., 1989, "Historical Development of Photogrammetric Methods and Instruments", American Society for Photogrammetry and Remote Sensing, 1989

Boochs F., 1992, "Object Space Based Correlation with Additional Information", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 127 - 132.

Boochs F., 1987, "Off-Line Compilation of Photogrammetric Stereo Models Using Digital Image Correlation", Photogrammetria, 41 (1987), pp. 183 - 199

Boochs F., 1986, "Determination of A Digital Terrain Model from Space Metric Camera Pictures by Means of Digital Image Correlation", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 79 - 92.

Castleman K., 1979, "Digital Image Processing", Prentice Hall, Englewood Cliffs, N.J.

Cho W. S., Schenk T., 1992, "Resampling Digital Imagery to Epipolar Geometry", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 404 - 408.

Claus M., 1984, "Digital Terrain Models through Digital Stereo Correlation", Photogrammetria, 39 (1984), pp. 183 - 192.

Church E., 1942 - 1948, "A Series of Bulletins on Aerial Photogrammetry", No. 13 - No. 19, Syracuse University.

Day T., 1989, "DEM Production by Stereo-matching SPOT Image-pairs: A Comparison of Algorithms", Image and Vision Computing, 7 (2), pp. 95 - 101.

Diehl H., Heipke C., 1992, "Surface Reconstruction from Data of Digital Line Cameras by Means of Object Based Matching", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 287 - 294.

Ding M. Y., 1992, "The Evaluation of Acquisition Probability in Image Matching", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 375 - 378.

Dowman L. J., 1984, "Problems and Some Solutions in Digital Correlation for Photogrammetric Profiling", Photogrammetria, 39 (1984), pp. 155 - 167.

Doyle F. J., 1964, "The historical Development of Analytical Photogrammetry", Photogrammetric Engineering, Vol. 30, pp. 259 - 265

Ebner H., Fritsch D., 1986, "The Multigrid Method and Its Application in Photogrammetry", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 141 - 149.

Ehlers M., 1984, On the Probability of Genealized Correlation Functions", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 25, Part A 3a, Commission III, pp. 139 - 150.

Ehlers M., Welch R., 1987, "Stereo Correlation of Landsat TM Image", Photogrammetric Engineering and Remote Sensing, Vol. 53, No. 9, pp. 1231 - 1237.

Ehlers M.., 1985, "The Effects of Image Noise on Digital Correlation Probability", Photogrammetric Engineering and Remote Sensing, Vol. 51, No. 3, pp. 357 - 365.

Ehlers M., 1982, "Increase in Correlation Accuracy of Remote Sensing Imagery by Digital Filtering", Photogrammetic Engineering and Remote Sensing, Vol. 48, No. 3, pp. 415 - 420.

El-Hakim S. F., 1989, "A Hierarchical Approach to Stereo Vision", Photogrammetric Engineering and Remote Sensing, Vol. 55, No. 4, pp. 443 - 448.

ERIO Technologies, 1991, "MAPP/PAL/ALBANY User's Manual", 800 Sunrise Avenue, Suite C, Roseville, CA 95661, Telephone: 9160782-6671.

Faintich M. B., 1984, "State-of-the-art and Future Needs for Development of Digital Terrain Models", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 25, Part A 3a, Commission III, pp. 180 - 196.

Fallvik J. O., 1986, " Image Restoration Methods as Preprocessing Tools in Digital Stereo Matching", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 223 - 235.

Förstner W., Gülch E., 1987, "A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features", ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, June, 1987, pp. 281 - 305.

Förstner W., 1986, "A Feature Based Correspondence Algorithm for Image Matching", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 150 - 163.

Förstner W., 1984, "Quality Assessment of Object Location and Point Transfer Using Digital Image Correlation Techniques", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 25, Part A 3a, Commission III, pp. 197 - 219.

Förstner W., 1982, "On the Geometric Precision of Digital Correlation", International Archives of Photogrammetry and Remote Sensing, Symposium Helsinki, Commission III,, Vol. 24, Part 3, pp. 176 - 189.

Frederiksen P., Jacobi O., 1986, "Optimal Sample Spacing in Digital Elevation Models", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 252 - 259.

Frederiksen P., Jacovi O., Kubik K., 1985, "A Review of Current Trade in Terrain Modeling", ITC Journal, 1985-2, pp. 101 - 106.

Fukushima Y., 1988, "Generation of DTM using SPOT Image Near Mt. Fuji by Digital Image Correlation", International Archives of Photogrammetry and Remote Sensing, Vol. 27, Part B3, Commission III, pp. 225 - 234.

Ghaffary B. K., 1985, "A Review of Image Techniques", SPIE 596, Cannes, pp. 164 - 172.

Ghosh K. Sanjib, 1988, "Analytical Photogrammetry", Second Edition, Pergamon Press, 1988.

Gremain F., Skordas T., 1992, "A Computer Vision Method for Motion Detection Using Cooperative Kalman Filters", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 303 - 306.

Goung L., Zheng T., 1992, "Stereo Matching Using Artificial Neural Networks", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 417 - 421.

Greenfeld J. S., 1990, "An Automatic Relative Orientation Based on Computer Vision Principles", Proceedings of 1990 ACSM-ASPRS Annual Convention, Vol. 5, pp. 34 - 43.

Greenfeld J. S., 1988, "Stereo Matching without An Operator's Intervention", Proceedings of 1988 ACSM-ASPRS Annual Convention, Vol. 3, pp. 56 - 65.

Greenfeld J. S., Schenk A. F., 1989, "Experiments with Edge-Based Stereo Matching", Photogrammetric Engineering and Remote Sensing, Vol. 55, No. 12, pp. 1771 - 1777.

Gruen A. W., Baltsavias E. P., 1987, "Geometrical Constrained Mutliphoto matching", ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, June, 1987, pp. 204 - 231.

Gruen A. W., Baltsavias E. P., 1986, "High Precision Image Matching for Digital Terrain Model Generation", ISPRS, Vol. 26, Commission III, pp. 284 - 296.

Gülch E., 1991, "Results of Test on Image Matching of ISPRS WG III/4", ISPRS, Vol. 46, pp. 1 - 18.

Hahn M., 1992, "Towards Automatic DTM Verification Exploiting Stereo Orthphotos", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, 233 - 240.

Hannah M. J., 1989, "A System for Digital Stereo Image Matching", Photogrammetric Engineering and Remote Sensing, Vol. 55, No. 12, pp. 1765 - 1770.

Hassan M. M., 1986, "A Special Analysis Method for Estimating the Sampling Density of Digital Elevation Models", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 306 - 316.

Hattori S., Mori C., Uchida O., 1986, "A Coarse-to-Fine Correlation Algorithm Considering Occlusions", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 317 -328.

He G. P., Novak K., 1992, "Automatic Analysis of Highway Features from Digital Stereo Images", International Society of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 119 - 124.

Heipke C., 1992, "A Global Approach for Least-Squares Image Matching and Surface Reconstruction in Object Space", Photogrammetric Engineering and Remote Sensing, Vol. 58, No. 3, pp. 317 - 323.

HELAVA Associates Inc., 1996, "Socet Set System Administrators Manual", GDE Systems Inc. Telephone: 800-227-3104.

Helava U. V., 1992a, "State of the Art in Digital Photogrammetric Workstations", ASPRS/ACSM/RT 92 Technical Papers, Washington D. C., Vol. 2, pp. 10 - 18.

Helava U. V., 1992b, "Prospects in Digital Photogrammetry", ASPRS/ACSM/RT 92 Technical Papers, Washington D. C., Vol. 2, pp. 19 - 24.

Helava U. V., 1988, "Object Space Least Squares Correlation", Proceedings of 1988 ACSM-ASPRS Annual Convention, Vol. 3, pp. 46 - 55.

Helava U. V., 1987, "Digital Comparator Correlation System", ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, June, 1987, pp. 404 - 417.

Helava U. V., 1978, "Digital Correlation in Photogrammetric Instruments", Photogrammetria, Vol. 34, pp. 19 - 41.

Helava U. V., 1976, "Digital Correlation in Photogrammetric Instruments" International Archives of Photogrammetry and Remote Sensing, Congress Helsinki, Commission II, Vol. 21.

Helava U. V., 1966, "A Fast Automatic Plotter", Photogrammetric Engineering, Vol. 32, No. 1, pp. 58 - 66.

Helava U. V., 1960, "Analytical Plotter Using Incremental Computer", Photogrammetric Engineering, Vol. 26, No. 3, pp. 428 - 435.

Helava U. V., 1958, "Analytical Plotter in Photogrammetric Production Line", Photogrammetric Engineering, Vol. 24, No. 5, pp. 794 - 797.

Hellwich O., Faig W., 1992, "Graph-based Matching of Stereo Image Features", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 307 - 317.

Hobrough G. L., 1959, "Automatic Stereo Plotting", Photogrammetric Engineering, Vol. 25, No. 5, pp. 763 - 769.

Hoff W., 1989, "Surface from Stereo: Integrating Feature Matching, Disparity, Estimation and Contour Detection", IEEE-PAMI Vol. 11, No. 2, pp. 121 -136.

Huertas A., Medioni G., 1986, "Detection of Intensity Changes with Subpixel Accuracy Using Lapacian-Guassian Masks", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 5, pp. 651 -664.

Johnson E. C., 1961, "System Design of a Digital Control Computer for an Analytical Stereoplotter", Photogrammetric Engineering, Vol. 27, pp. 583 - 589.

Kaiser B., S. M., Wrobel B. P., 1992, "Application of Image Pyramid for Surface Reconstruction with FAST Vision (=Facets Stereo Vision)", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 341 - 374.

Kölbl O., Bach U., Gasior D., de laporte K., 1992, "Multi-Templet-Matching for the Automation of Photogrammetric Measurements", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 540 - 548.

Kölbl O., 1991, "The DSR 15T - A System for Automatic Image Correlation", 1991 ACSM-ISPRS Annual Convention, Vol. 15, pp. 218 - 227.

Kölbl O., 1990, "Automatic Derivation of A DTM", ASPRS-ACSM Annual Convention, Vol. 5, pp. 54 - 66.

Kölbl O., 1989, "Replacement of the Human Operator by Image Processing Shown on the Example of the Elaboration of A DTM", 1989 ASPRS-ACSM Annual Convention, Vol. 2, pp. 126 - 141.

Kölbl O., Boutaleb A. K., Penis C., 1987, "A Concept for Automatic Derivation of A Digital Terrain Model with the KERN DSR11", ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, June, 1987, pp. 306 - 317.

Kropatsch W. G., 1985, "A Pyramid That Grows by Powers of Two", Pattern Recognition Letters, Vol. 3, pp. 315 - 322.

Krupnik A., Schenk T., 1992, "Segmentation of Edges in 3-D Object Space", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 522 - 527.

Krzystek P., Ackermann F., 1995, "New Investigation into the Practical Performance of Automatic DEM Generation", ACSM/ASPRS Annual Convention & Exposition Technical Papers, Vol. 2, pp. 488 - 500.

Lanckton A. H., 1969, "Analytical Stereoplotter Development", Photogrammetric Engineering, Vol. 35, No. 11, pp. 1160 - 1168.

Lawrence C. H., 1967, "Stereomat IV, Automatic Plotter", Photogrammetric Engineering, Vol. 33, No. 4, pp. 394 - 406.

Lee D. C., Schenk T., 1992, "Image Segmentation from Texture Measurement", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 195 - 199.

Li D. R., Shao J. L., 1994, "The Wavelet and Its Application in Image Edge Detection", ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 49, No. 3, pp. 4 - 11.

Li M. X., 1990, "High-precision Relative Orientation Using Feature-based Matching Techniques", ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 44, pp. 311 - 324.

Light D. L., 1966, "The Orientation Matrix", Photogrammetric Engineering, Vol. 32, pp. 434 - 438.

Lin Z. J., 1986, "Multi-Criterions for Similarity Assessment in Photogrammetric Image Correlation", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 415 - 426.

Lo K. C., Mulder N. J., 1992, "High Precision DEM Generation from SPOT Stereo Imagery by Object Space Least Square Matching", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 133 - 137.

Luhman T., Altrogge G., 1986, "Inerest-Operator for Image Matching", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 459 - 474.

Makarovic B., 1992, "Consideration on Image Matching - An Engineering Perspective", ISPRS, Vol. 29, Part 2, pp. 613 - 622.

Masry S. E., 1974, "Digital Correlation Principles", Photogrammetric Engineering, Vol. 40, No. 3, pp. 303 - 308.

Medioni G., Nevatia R., 1984, "Matching Images Using Linear Features", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 6, pp. 675 - 685.

Miller S. B., 1992a, "Automatic Elevation Extraction and the Digital Photogrammetric Workstation", ASPRS-ACSM Annual Convention, pp. 572 - 580.

Miller S. B., 1992b, "Softcopy Photogrammetric Workstations", Photogrammetric Engineering and Remote Sensing, Vol. 58, No. 1, pp. 77 - 83.

Pertl A., 1985, "Digital Image Correlation with An Analytical Plotter", Photogrammetria, 40(1985), pp. 9 - 19.

Piechel J., 1986, "Investigations of Different Interest Operators for DTM-Generation by Epipolar Line Correlation", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 564 - 572.

Pratt W., 1978, "Digital Image Processing", John Wiley and Sons, New York, N.Y.

Qiu Z. C., Liu Y. T., 1992, "A Research of Boundary Extraction Based on Zero Crossing of Second Directional Derivatives", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 43 - 46.

Rauhala U. A., 1992, "Automation of Global Digital Image Mapping by Array Algebra", ASPRS/ACSM/RT Technical Papers, Washington, D. C., Vol. 2, pp. 65 - 75.

Rauhala U. A., 1987, "Fast Compiler Positioning Algorithms and Techniques of Array Algebra in Analytical and Digital Photogrammetry", ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, June, 1987, pp. 156 - 178.

Reis M. L., 1992, "Optimized Algorithm for Three Dimensional Object Reconstruction", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 335 - 340.

Rosenholm D., 1987, "Least Squares Matching Method: Some Experimental Results", Photogrammetric Records, 12 (70), pp. 493 - 512.

Rosenholm D., 1987, "Empirical Investigation of Optimal Window Size Using the Least Squares Image Matching Method", Photogrammetria (PRS), Vol. 42, pp. 113 - 125.

Rosenholm D., 1987, "Multi-point Matching Using the Least Squares Technique for Evaluation of Three-Dimensional Model", Photogrammetric Engineering and Remote Sensing, Vol. 27, pp. 621 - 626.

Rosenholm D., 1986, "Accuracy Improvement of Digital Matching for Evaluation of Digital Terrain Models", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 573 - 587.

Russell R. L., McClain R. A., Landell B. P., 1992, " A 3-D Model Extraction System", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 446 - 451.

Saleh R. A., Scarpace F. L., 1992, "Multispectral Matching techniques for DTM Generation", ASPRS/ACSM/RT Technical Papers, Washington D. C., Vol. 2, pp. 36 - 46.

Scarano F. A., 1976, "A Digital Elevation Data Collection System", Photogrammetric Engineering and Remote Sensing", Vol. 42, No. 4, pp. 489 - 496.

Schenk A., 1986, "Stereo Matching Using Line Segments of Zero Crossings", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 602 - 607.

Schenk T., Toth C. K., 1992, "Reconstructing Small Surface from Multiple Images", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 255 - 258.

Schewe H., Förstner, 1986, "The Program PALM for Automatic Line and Surface Measurement Using Image Matching Techniques", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 608 - 622.

Schewe H., Förstner W., 1986, "A Feature Based Correspondence Algorithm for Image Matching", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 150 - 163.

Schneider F., Hahan M., 1992, "Matching Images of Different Geometric Scale", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 295 -302.

Smith J. A., Lin T. L., Ranson K. J., 1980, "The Lambertian Assumption and Landsat Data", Photogrammetric Engineering and Remote Sensing, Vol. 46, No. 9, pp. 1183 - 1189.

Stefanidis A., Schenk T., 1992, "On the Application of Scale Space Techniques in Digital Photogrammetry", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 586 - 590.

Stokes John, 1986, "Image Matching with Phase Shift Methods", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 26, Commission III, pp. 638 - 652.

Toth Charles K., Schenk T., 1992, "On Matching Image Patches under Various Geometrical Constraints", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 400 - 403.

Trinder J. C., 1975, "Autocorrelation Function in Stereoscopy", Photogrammetric Engineering and Remote Sensing, Vol. 41, No. 3, pp. 325 - 330.

Vision International, a division of Autometric Inc., 1996, "SoftPlotter Product Functional Description, Version 1.6", 502 Earth City Expressway, Suite 203, Earth City, MO 63045. Telephone: 314-770-2015.

Wang S., Haralick R. M., Campbell J., 1984, "Relative Elevation Determination from LANDSAT Imagery", Photogrammetria, 39(1984), pp. 193 - 215.

Wang Z. Z., 1984, "Principles of Photogrammetry", Surveying Press, People's Republic of China, 1984.

Wrobel B. P., 1991, "Least-square methods for surface reconstruction from images", ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 46(1991), pp. 67 - 84.

Wrobel B. P., 1987, "Facet Stereo vision(FAST Vision) - A New Approach to Computer Stereo Vision and to Digital Photogrammetry", Proceedings of Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, pp. 231 - 258.

Zhang J. Q., Zhang Z. X., Wang Z. H., 1992, " High Accurate Location on Digital Image and Application in Automatic Relative Orientation", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 78 - 82.

Zhang J. Q., Fan Q. S., 1992, "Image Analysis Based on Mathematical Morphology", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, pp. 72 - 77.

Zhang Y. N., 1992, "A New Stereo Matching Approach in Image/Object Dual Space", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 534 - 539.

Zhang Z. X., Min Y. R., Zhang J. Q., 1992, "Image Segmentation Based on Hough Transformation", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 633 - 638.

Zong J., Li J. C., Schenk T., 1992, "Aerial Image Matching Based on Zero-Crossings", ISPRS International Archives of Photogrammetry and Remote Sensing, Vol. 28, Commission III, pp. 144 - 150.

ACKNOWLEDGMENTS

I would like to thank Dr. Kandiah Jeyapalan, my major professor, for his guidance, assistance and supervision throughout my research project. I could not have finished this research without his help. I also wish to thanks Dr. U. Sunday Tim, Dr. Mushtaq-ur Rahman, Dr. Dah-Yinn Lee, Dr. Thomas Colvin, and Dr. John M. Pitt for their input and for serving on the program of study committee.

I also would like to thank Civil and Construction Engineering Department for the teaching assistant (TA) position, and College of Engineering and Iowa State University for the research assistant (RA) position which supported this study during the four years it took to complete.

I also wish to thank Mrs. Alice Welch from Iowa Department of Transportation for her support and help.

Software packages from Vision International, HELAVA Inc. and ERIO were used in this research. I want to express my appreciation to them. I am also indebted Dr. Jeyapalan for his software which was used in this research.

A special thanks to my fellow graduate students, Mr. Sarath, Mr. Ruifeng, Mr. Rocky, and Mr. Kang, who assisted my in various aspects of this research.

Finally, I wish to express my gratitude to my wife, Jing Zhang, for her support and encouragement when I needed it.